

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 5, 2004, 05:30:38 ; Search time 14364 Seconds
(without alignments)
11516.247 Million cell updates/sec

Title: US-09-679-643-1
Perfect score: 3498
Sequence: 1 ggtaccataaataacacatc.....gaagagtataagaagaagctt 3498

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3498	100.0	3498	6	AR096793 Sequence
2	3498	100.0	3498	6	AR125115 Sequence
3	3487	99.7	3499	1	SACFG
4	3416.2	97.7	37669	1	AC074317
5	3399.6	97.2	4709	6	AR354469 Sequence
6	3277	93.7	110000	1	Continuation (9 of
7	3245	92.8	290150	1	AP004824
8	2960.6	84.6	348477	1	AP003360
9	2853.4	81.6	299050	1	AP003131
10	2797.4	80.0	2799	6	AX583664
11	2746.6	78.5	2781	6	AX619958
12	2686.8	76.8	110000	1	Continuation (9 of
13	1776.4	50.8	2792	6	AX343875
14	1776.4	50.8	2792	6	AX768506
15	917.8	26.2	2721	1	AF404823
16	619.2	17.7	5202	1	AF245041
17	619.2	17.7	5406	6	BD245129
18	619.2	17.7	5406	6	AR410982
19	619.2	17.7	5406	6	AR482275

C	20	592.8	16.9	90802	1	AE016752	Staphyloc
	21	589.6	16.9	4627	1	AF259905	Staphyloc
	22	589.6	16.9	4627	6	AR485859	Sequence
	23	589.6	16.9	4627	6	AX145223	Sequence
	24	584.8	16.7	1485	6	AR483599	Sequence
	25	584.8	16.7	1485	6	AX141313	Sequence
	26	583.4	16.7	4656	1	AY048593	Staphyloc
C	27	574.4	16.4	117662	1	AP004831	Staphyloc
	28	568.6	16.3	2969	6	AR453524	Sequence
	29	568.6	16.3	2969	6	BD140510	Staphyloc
	30	568.6	16.3	2971	1	SAA224764	Fibrinoge
	31	568.4	16.2	2133	1	AY510088	Staphyloc
	32	568	16.2	2739	6	AX583634	Sequence
	33	562	16.1	269050	1	AL935255	Lactobaci
	34	558.2	16.0	1059	6	AX821821	Sequence
C	35	555.6	15.9	110000	1	BX571857_26	Continuation (27 o
C	36	541.8	15.5	99802	1	BX571857_27	Continuation (28 o
	37	525.4	15.0	2151	6	AX583629	Sequence
	38	510	14.6	2631	6	AX620920	Sequence
C	39	510	14.6	3007	6	AX768448	Sequence
	40	510	14.6	11050	6	AR353978	Sequence
C	41	510	14.6	49898	1	AC027136	Staphyloc
C	42	502.6	14.4	2582	6	AX768449	Sequence
C	43	502.6	14.4	2582	6	AX768519	Sequence
C	44	502	14.4	109528	1	AP003366	Staphyloc
C	45	502	14.4	117016	1	AP003138	Staphyloc

ALIGNMENTS

RESULT 1
AR096793
LOCUS AR096793 3498 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 1 from patent US 6008341.
ACCESSION AR096793
VERSION AR096793.1 GI:10025907
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3498)
AUTHORS Foster,T.James, and McDevitt,D.Leo.
TITLE S. aureus fibrinogen binding protein gene
JOURNAL Patent: US 6008341-A 1 28-DEC-1999;
FEATURES
Location/Qualifiers
source 1..3498
/organism="unknown"
/mol_type="unassigned DNA"

Query Match	100.0%	Score	3498	DB	6	Length	3498
Best Local Similarity	100.0%	Pred. No.	0				
Matches	3498	Conservative	0	Mismatches	0	Indels	0
						Gaps	0
QY	1	GGTACCATTAATACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTCATTA	60				
Db	1	GGTACCATTAATACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTCATTA	60				
QY	61	GGTAGAGTTTCATATTAAATAATAAATAAATGTTTGCAATCAATCGTACGTTGCTGTTTGA	120				
Db	61	GGTAGAGTTTCATATTAAATAATAAATAAATGTTTGCAATCAATCGTACGTTGCTGTTTGA	120				
QY	121	ATTCCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA	180				
Db	121	ATTCCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA	180				
QY	181	TATCGATACAAATTAATTCCTATATGCAATTTTAGTATTAATTCATTAACAGAGATT	240				
Db	181	TATCGATACAAATTAATTCCTATATGCAATTTTAGTATTAATTCATTAACAGAGATT	240				
QY	241	AAATATATCTTAAAGGGTATATAGTTATATATAAATGACCTTTTAAAGAGGGATATAA	300				

applicants

Db 241 AAATAATCTTTAAAGGTATATAGTTAATAATAAATGACTTTTAAAAAGAGGGAATAAA 300
QY 301 ATGNAATGAAGAAAAAGAAAAACAGCAATTCGGAAAAAATCGAATTCGGTGGCTTCA 360
Db 301 ATGAATATGAAGAAAAAGAAAAACAGCAATTCGGAAAAAATCGAATTCGGTGGCTTCA 360
QY 361 GTGCTGTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAGT 420
Db 361 GTGCTGTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAGT 420
QY 421 GAAATATGTTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAGT 480
Db 421 GAAATATGTTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAGT 480
QY 481 AGCGTTAGTGTCACTCACTAAACAGACACACAAACGTTGAGTGATCTAAACATCGTCA 540
Db 481 AGCGTTAGTGTCACTCACTAAACAGACACACAAACGTTGAGTGATCTAAACATCGTCA 540
QY 541 AACACTAATATGCGGAAACGAGTGTGGCGCAAAATCCAGCACACAGGAAACGACACAA 600
Db 541 AACACTAATATGCGGAAACGAGTGTGGCGCAAAATCCAGCACACAGGAAACGACACAA 600
QY 601 TCATCATCAACAAATGCAACTACCGGCAAAACGCGGTAACTGGTGAAGCTACTACTACG 660
Db 601 TCATCATCAACAAATGCAACTACCGGCAAAACGCGGTAACTGGTGAAGCTACTACTACG 660
QY 661 ACACGATCAAGCTAATACACCGGCAACACTCAATCAAGCAATACAAATGCGGAGAA 720
Db 661 ACACGATCAAGCTAATACACCGGCAACACTCAATCAAGCAATACAAATGCGGAGAA 720
QY 721 TTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATCTAATACAGTATCATCTGTA 780
Db 721 TTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATCTAATACAGTATCATCTGTA 780
QY 781 AATTCACTCAAAATCTTACAAATGCGGAAATGTTTCAACACGCAAGTACTTCAACT 840
Db 781 AATTCACTCAAAATCTTACAAATGCGGAAATGTTTCAACACGCAAGTACTTCAACT 840
QY 841 GAAGCAACACTTCAACAAATGAATCAGTCCAGAGTACAGATGCAAGTAATAAGAT 900
Db 841 GAAGCAACACTTCAACAAATGAATCAGTCCAGAGTACAGATGCAAGTAATAAGAT 900
QY 901 GTAGTTAATCAAGCGGTTAATACAGTGCCTGAGATGAGAGCATTTAGTTAGCGCA 960
Db 901 GTAGTTAATCAAGCGGTTAATACAGTGCCTGAGATGAGAGCATTTAGTTAGCGCA 960
QY 961 GTAGCTGAGATGACCGGAGCTGGGACAGATATACGAATCAGTTGACGAAATGTGACA 1020
Db 961 GTAGCTGAGATGACCGGAGCTGGGACAGATATACGAATCAGTTGACGAAATGTGACA 1020
QY 1021 GTTGTATTGACTCTGTGACGACTGTGTATCGCACCAAGCAGGTTATGTCAAACTGAAT 1080
Db 1021 GTTGTATTGACTCTGTGACGACTGTGTATCGCACCAAGCAGGTTATGTCAAACTGAAT 1080
QY 1081 TATGTTTATTCAGTGCCTTAATCTCTGTTTAAAGGTGACATTCAAAATAAATCTACTCT 1140
Db 1081 TATGTTTATTCAGTGCCTTAATCTCTGTTTAAAGGTGACATTCAAAATAAATCTACTCT 1140
QY 1141 AAAGAAATTAACCTTAAATGGTGTAACTTCACTGCTAAAGTGCCACCAATTTATGGCTGA 1200
Db 1141 AAAGAAATTAACCTTAAATGGTGTAACTTCACTGCTAAAGTGCCACCAATTTATGGCTGA 1200
QY 1201 GATCAAGTATTGGCAAAATGGTGTAAATCGATAGTGTGTTATTTATACATTTACA 1260
Db 1201 GATCAAGTATTGGCAAAATGGTGTAAATCGATAGTGTGTTATTTATACATTTACA 1260
QY 1261 GACTATGTAATTAACAGATGATAAAGCAACTTTGACCAATGCCCGCTTATATTGAC 1320
Db 1261 GACTATGTAATTAACAGATGATAAAGCAACTTTGACCAATGCCCGCTTATATTGAC 1320
QY 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAACA 1380
Db 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAACA 1380

QY 1381 GCAACAAAAACAGTATTAGTAGATTATGAAAAATATCGTAAAGTTTATAAATTATCTATT 1440
Db 1381 GCAACAAAAACAGTATTAGTAGATTATGAAAAATATCGTAAAGTTTATAAATTATCTATT 1440
QY 1441 AAAGGTACAATTCGACCAATTCGATAAAACAAATATATACGTATTCGTACAGCAATTTATGTC 1500
Db 1441 AAAGGTACAATTCGACCAATTCGATAAAACAAATATATACGTATTCGTACAGCAATTTATGTC 1500
QY 1501 AATCCAGTGCAGATACGTTATTTCGCGCGTTTAAACAGTAAATTTAAACACCAAAATACG 1560
Db 1501 AATCCAGTGCAGATACGTTATTTCGCGCGTTTAAACAGTAAATTTAAACACCAAAATACG 1560
QY 1561 GATAGTAAATGATTAATAGATCAGCAAAATAACAAGTATTTAAAGTATATAAAGTAGATAAT 1620
Db 1561 GATAGTAAATGATTAATAGATCAGCAAAATAACAAGTATTTAAAGTATATAAAGTAGATAAT 1620
QY 1621 GCAGCTGATTATCTGAAAGTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCACTAAT 1680
Db 1621 GCAGCTGATTATCTGAAAGTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCACTAAT 1680
QY 1681 AGTGTGAATATTACATTCCTCAATCCAAATCAATATAAAGTAGAGTTTAAATACCGCTGAT 1740
Db 1681 AGTGTGAATATTACATTCCTCAATCCAAATCAATATAAAGTAGAGTTTAAATACCGCTGAT 1740
QY 1741 GATCAAAATTACAAACACCGTATATAGTAGTTTAAATGTCATATTTGAATCCGAAATAGCAAA 1800
Db 1741 GATCAAAATTACAAACACCGTATATAGTAGTTTAAATGTCATATTTGAATCCGAAATAGCAAA 1800
QY 1801 GGTGATTAGCTTTACGTTCACTTTATATGGGTATTAACCTCGAATATAAATTTGGGCTCT 1860
Db 1801 GGTGATTAGCTTTACGTTCACTTTATATGGGTATTAACCTCGAATATAAATTTGGGCTCT 1860
QY 1861 ATGTCATGGGACACAGTAGCAATTTAAACCGATCAGGTTCTGTGTGACGCTATCGAT 1920
Db 1861 ATGTCATGGGACACAGTAGCAATTTAAACCGATCAGGTTCTGTGTGACGCTATCGAT 1920
QY 1921 AAACAGTGTGTTCTGTAAACCTGATGAGCTTGGTGAATTTGAACCAATTTCCAGAGGAT 1980
Db 1921 AAACAGTGTGTTCTGTAAACCTGATGAGCTTGGTGAATTTGAACCAATTTCCAGAGGAT 1980
QY 1981 TCAGATTCTGACCAAGGTTAGATTCTGGGACGCAATTTCAATTCAGATAGCGGTTTCAGAT 2040
Db 1981 TCAGATTCTGACCAAGGTTTCAGATTCTGGGACGCAATTTCAATTCAGATAGCGGTTTCAGAT 2040
QY 2041 TCGGTTAGTGAATCTTACATCAGATAGTGGTTTCAGATTTCAGGAGTGAATTCAGATTTCAGCA 2100
Db 2041 TCGGTTAGTGAATCTTACATCAGATAGTGGTTTCAGATTTCAGGAGTGAATTCAGATTTCAGCA 2100
QY 2101 AGTGATTTCAGACTCAGCGAGTGAATTCAGATTTCAGCAAGCGAATTCGACTCAGCGAGCGAT 2160
Db 2101 AGTGATTTCAGACTCAGCGAGTGAATTCAGATTTCAGCAAGCGAATTCGACTCAGCGAGCGAT 2160
QY 2161 TCCGACTCAGCAATGACTCCGATTTCAGATAGCGAATTCGACTCAGACAGTGAATTCAGAT 2220
Db 2161 TCCGACTCAGCAATGACTCCGATTTCAGATAGCGAATTCGACTCAGACAGTGAATTCAGAT 2220
QY 2221 TCCGACAGTGAATTCAGATTTCAGATAGCGAATTCGACTCAGACAGTGAATTCAGATTTCAGAT 2280
Db 2221 TCCGACAGTGAATTCAGATTTCAGATAGCGAATTCGACTCAGACAGTGAATTCAGATTTCAGAT 2280
QY 2281 AGCGATTTCAGATTTCAGATAGCGAATTCAGATTTCGCAAGTGAATTCGACTCAGACAGCGAT 2340
Db 2281 AGCGATTTCAGATTTCAGATAGCGAATTCAGATTTCGCAAGTGAATTCGACTCAGACAGCGAT 2340
QY 2341 TCTGACTCCGACAGTGAATTCGACTCAGACAGCGAATTCAGATTTCGCAAGTGAATTCGCGAC 2400
Db 2341 TCTGACTCCGACAGTGAATTCGACTCAGACAGCGAATTCAGATTTCGCAAGTGAATTCGCGAC 2400
QY 2401 TCAGATAGCGATTTCGACTCAGATAGCGAATTCAGATTTCAGACAGCGAATTCAGATTTCAGAC 2460
Db 2401 TCAGATAGCGATTTCGACTCAGATAGCGAATTCAGATTTCAGACAGCGAATTCAGATTTCAGAC 2460

2461 AGCGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2520
2461 AGCGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2520
2521 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2580
2521 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2580
2581 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2640
2581 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2640
2641 AGTGAATTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2700
2641 AGTGAATTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2700
2701 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2760
2701 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2760
2761 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2820
2761 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2820
2821 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2880
2821 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2880
2881 AGCGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2940
2881 AGCGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 2940
2941 ACTAATGCTTCTAATAAATGAGTTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3000
2941 ACTAATGCTTCTAATAAATGAGTTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3000
3001 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3060
3001 TCAGATTGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3060
3061 CTACTTTTCAGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3120
3061 CTACTTTTCAGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3120
3121 TAATCATATGATTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3180
3121 TAATCATATGATTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3180
3181 AATATATGCTTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3240
3181 AATATATGCTTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3240
3241 TAGATGCTTATATATTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3300
3241 TAGATGCTTATATATTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3300
3301 GAAATATGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3360
3301 GAAATATGAGTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3360
3361 ATGTGTATCACAATTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3420
3361 ATGTGTATCACAATTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3420
3421 TCCATATGCTTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3480
3421 TCCATATGCTTTCAGATGACGATTCAGATTCGACAGTGAATCAGATTCGACAGTGAAC 3480
3481 AGAGTATTAAGAAAGCTT 3498
3481 AGAGTATTAAGAAAGCTT 3498

RESULT 2
AR125115
LOCUS
DEFINITION
SEQUENCE 1 from patent US 6177084.
AR125115
VERSION
AR125115.1 GI:14111177
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 3498)
AUTHORS
Foster, T. James, and McDevitt, D. Leo.
TITLE
S. aureus fibrinogen binding protein
JOURNAL
Patent: US 6177084-A 1 23-JAN-2001;
FEATURES
Location/Qualifiers
source
1. 3498
/organism="unknown"
/mol_type="unassigned DNA"

exp

Query Match 100.0%; Score 3498; DB 6; Length 3498;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTAACATATAATACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATTCATTA 60
DB 1 GGTAACATATAATACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATTCATTA 60
QY 61 GGTAACATATAATACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATTCATTA 120
DB 61 GGTAACATATAATACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATTCATTA 120
QY 121 ATCTTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTTTGGATTAATAAAA 180
DB 121 ATCTTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTTTGGATTAATAAAA 180
QY 181 TATCGATACAAATTAATTCGTATTAATGCAATCAATCGTACGTTGCTGTTGTA 240
DB 181 TATCGATACAAATTAATTCGTATTAATGCAATCAATCGTACGTTGCTGTTGTA 240
QY 241 AATATATCTTAAAGGGTATAGTATTAATAATAAATGACTTTTAAAGAGGGAATAA 300
DB 241 AATATATCTTAAAGGGTATAGTATTAATAATAAATGACTTTTAAAGAGGGAATAA 300
QY 301 ATGAATATGAAGAAAGAAAGAAACACGCAATTCGGAAAAATTCGATTCGGGTTC 360
DB 301 ATGAATATGAAGAAAGAAAGAAACACGCAATTCGGAAAAATTCGATTCGGGTTC 360
QY 361 GTGCTTGTAGGTACGTTAATTCGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAGT 420
DB 361 GTGCTTGTAGGTACGTTAATTCGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAGT 420
QY 421 GAAATATAGTTCACCAATCTCATAGCGCAAGTAAAGCAAGTAAATGATTCAGT 480
DB 421 GAAATATAGTTCACCAATCTCATAGCGCAAGTAAAGCAAGTAAATGATTCAGT 480
QY 481 AGCGTTAGTTCGACCTTAAACACAGACACAAACGTCAGTAAAGAGCAGATGCAAGT 540
DB 481 AGCGTTAGTTCGACCTTAAACACAGACACAAACGTCAGTAAAGAGCAGATGCAAGT 540
QY 541 AACACTAATAATGGGAAACGAGTGTGGGCAAAATCCAGCAACACAGAAACGACAA 600
DB 541 AACACTAATAATGGGAAACGAGTGTGGGCAAAATCCAGCAACACAGAAACGACAA 600
QY 601 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTACG 660
DB 601 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTACG 660
QY 661 ACAACGAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA 720
DB 661 ACAACGAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA 720
QY 721 TTATGGAATCAACCAAGTAAATGAAACGACTTTTATGATACTTAATACAGTATCATCTGTA 780

Db 601 ATCATCATCAACAAATGCAACTACGGAAGAAACGCCGGTAACTGGTGAAGCTACTACTAC 660
Qy 660 GACAAACGAATCAAGCTTAATACACCGCAACAACTCAATCAAGCAATCAAAATCGGAGGA 719
Db 661 GACAAACGAATCAAGCTTAATACACCGCAACAACTCAATCAAGCAATCAAAATCGGAGGA 720
Qy 720 ATTAGTGAATCAAAACAAAGTAAATGAACGACATTTTAAATGATACTAATACAGTATCATCTGT 779
Db 721 ATTAGTGAATCAAAACAAAGTAAATGAACGACATTTTAAATGATACTAATACAGTATCATCTGT 780
Qy 780 AAATTCACCTCAAAATTTCTACAAATGCGGAAATATGTTTCAACAAACGCAAGATCTTCAAC 839
Db 781 AAATTCACCTCAAAATTTCTACAAATGCGGAAATATGTTTCAACAAACGCAAGATCTTCAAC 840
Qy 840 TGAAGCAACCTTTCAACAAATGAATCAGCTCCACAGATACAGATGCAAGATTAATAAGA 899
Db 841 TGAAGCAACCTTTCAACAAATGAATCAGCTCCACAGATACAGATGCAAGATTAATAAGA 900
Qy 900 TGTAGTTAATCAAGCGGTTAATACAAGTGGCGCTAGAAATGAGACATTTAGTTAGCGGC 959
Db 901 TGTAGTTAATCAAGCGGTTAATACAAGTGGCGCTAGAAATGAGACATTTAGTTAGCGGC 960
Qy 960 AGTAGCTGCAGATGCAACCGGACGCTGGCAAGATATACGAATCAGTTGACGAATGTGAC 1019
Db 961 AGTAGCTGCAGATGCAACCGGACGCTGGCAAGATATACGAATCAGTTGACGAATGTGAC 1020
Qy 1020 AGTTGGTATTCAGCTCTGTGACGCTGTGTATCCGCAACAGCAGGTTATGTCAAACTGAA 1079
Db 1021 AGTTGGTATTCAGCTCTGTGACGCTGTGTATCCGCAACAGCAGGTTATGTCAAACTGAA 1080
Qy 1080 TTATGGTTTTTTCAGTGCCCTAATTCGCTGTTTAAAGGTGACACATTCAAAATTAACCTGACC 1139
Db 1081 TTATGGTTTTTTCAGTGCCCTAATTCGCTGTTTAAAGGTGACACATTCAAAATTAACCTGACC 1140
Qy 1140 TAAAGAAATTAACCTTAATGCTGTAACTTCACTGCTTAAGTGCACCAATTAATGGCTGG 1199
Db 1141 TAAAGAAATTAACCTTAATGCTGTAACTTCACTGCTTAAGTGCACCAATTAATGGCTGG 1200
Qy 1200 AGATCAAGTATGCAAAATGCTGTAACTCGATAGTGAATGCTTAATTTATACATTTAC 1259
Db 1201 AGATCAAGTATGCAAAATGCTGTAACTCGATAGTGAATGCTTAATTTATACATTTAC 1260
Qy 1260 AGACTATGTAATACTAAAGATGATGTAAAGCAACTTTGACCAATGCGCGCTTATTTGA 1319
Db 1261 AGACTATGTAATACTAAAGATGATGTAAAGCAACTTTGACCAATGCGCGCTTATTTGA 1320
Qy 1320 CCCTGAAATGTTAAAGACAGGTAATGTGACATTTGCTACTGCGCATAGGTAGTACAAAC 1379
Db 1321 CCCTGAAATGTTAAAGACAGGTAATGTGACATTTGCTACTGCGCATAGGTAGTACAAAC 1380
Qy 1380 AGCAAAACAAACAGTATTAGTAGATTATGAAATAATGTTAACTTTATCTAT 1439
Db 1381 AGCAAAACAAACAGTATTAGTAGATTATGAAATAATGTTAACTTTATCTAT 1440
Qy 1440 TAAAGGTACAATTCGACCAAAATCGATATAAATAATACGTTATCGTCAGACAAATTTATGT 1499
Db 1441 TAAAGGTACAATTCGACCAAAATCGATATAAATAATACGTTATCGTCAGACAAATTTATGT 1500
Qy 1500 CAATCCAGTGGAGATACGTTATTTGCGCGGTTTAAACAGGTAATTTAAACCAAAATAC 1559
Db 1501 CAATCCAGTGGAGATACGTTATTTGCGCGGTTTAAACAGGTAATTTAAACCAAAATAC 1560
Qy 1560 GGATAGTAATGCAATTAATAGATACGAAATAACAGTATTAAAGTATATAAGTAGATAA 1619
Db 1561 GGATAGTAATGCAATTAATAGATACGAAATAACAGTATTAAAGTATATAAGTAGATAA 1620
Qy 1620 TGCAGCTGAATTTATCTGAAGTTACTTTGTGAATCCAGAAACTTTTGAGGATGTCACTAA 1679
Db 1621 TGCAGCTGAATTTATCTGAAGTTACTTTGTGAATCCAGAAACTTTTGAGGATGTCACTAA 1680
Qy 1680 TAGTGTGAATATTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATCCGCTGA 1739
Db 1681 TAGTGTGAATATTACATTTCCCAATCCCAATCAATATAAAGTAGAGTTTAAATCCGCTGA 1740

Qy 1740 TGATCAAAATTAACAAACCGTATATAGTAGTTGTTTAAATGGTCATATTTGATCCGAATAGCAA 1799
Db 1741 TGATCAAAATTAACAAACCGTATATAGTAGTTGTTTAAATGGTCATATTTGATCCGAATAGCAA 1800
Qy 1800 AGGTGATTTAGCTTTTACGTTCAACTTTATATGGGTAACTCGAATAATAATTTGGCGCTC 1859
Db 1801 AGGTGATTTAGCTTTTACGTTCAACTTTATATGGGTAACTCGAATAATAATTTGGCGCTC 1860
Qy 1860 TATGTCATGGGACAAACGAAGTAGCATTTAATAACGGATCAGGTTCTGGTACGGTATCGA 1919
Db 1861 TATGTCATGGGACAAACGAAGTAGCATTTAATAACGGATCAGGTTCTGGTACGGTATCGA 1920
Qy 1920 TAAACCAAGTTGTTCTCTGAAACAAACCTGATGAGCCCTGGTGAATTTGAACCAATTCAGAGGA 1979
Db 1921 TAAACCAAGTTGTTCTCTGAAACAAACCTGATGAGCCCTGGTGAATTTGAACCAATTCAGAGGA 1980
Qy 1980 TTCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTTCA 2039
Db 1981 TTCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTTCA 2040
Qy 2040 TTCGGGTAGTAGTTCTACATCAGATAGTGGTTTCAGATTTACGCGAGTGAATTCAGATTCAGC 2099
Db 2041 TTCGGGTAGTAGTTCTACATCAGATAGTGGTTTCAGATTTACGCGAGTGAATTCAGATTCAGC 2100
Qy 2100 AAGTGATTCAGACTCAGCGAGTGAATTCAGATTTACGCAAGCGGATTCGACTCAGCGAGCGA 2159
Db 2101 AAGTGATTCAGACTCAGCGAGTGAATTCAGATTTACGCAAGCGGATTCGACTCAGCGAGCGA 2160
Qy 2160 TTCGACTCAGACAAATGACTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGA 2219
Db 2161 TTCGACTCAGACAAATGACTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGA 2220
Qy 2220 TTCGCAAGTGCACACAGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGA 2279
Db 2221 TTCGCAAGTGCACACAGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGA 2280
Qy 2280 TAGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTGAATTCGACTCAGACAGCGA 2339
Db 2281 TAGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTGAATTCGACTCAGACAGCGA 2340
Qy 2340 TTTCTGACTCCGACAGTGAATTCGCACTCAGACAGGATTCAGATTCGCAAGTGAATTCGCA 2399
Db 2341 TTTCTGACTCCGACAGTGAATTCGCACTCAGACAGGATTCAGATTCGCAAGTGAATTCGCA 2400
Qy 2400 CTGAGATAGCGATTTCCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGA 2459
Db 2401 CTGAGATAGCGATTTCCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGA 2460
Qy 2460 CAGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTGAATTCGCAAGTGAATTCGCA 2519
Db 2461 CAGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTGAATTCGCAAGTGAATTCGCA 2520
Qy 2520 CTCGGATTCAGATAGCGATTCAGATTCGCAAGTGAATTCGCAAGTGAATTCGCAAGTGAATTC 2579
Db 2521 CTCGGATTCAGATAGCGATTCAGATTCGCAAGTGAATTCGCAAGTGAATTCGCAAGTGAATTC 2580
Qy 2580 CTCAGACAGTGAATTCGCAAGTGAATTCGCAAGTGAATTCGCAAGTGAATTCGCAAGTGAATTC 2639
Db 2581 CTCAGACAGTGAATTCGCAAGTGAATTCGCAAGTGAATTCGCAAGTGAATTCGCAAGTGAATTC 2640
Qy 2640 CAGTGACTCGGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 2699
Db 2641 CAGTGACTCGGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 2700
Qy 2700 TTCGACTCAGATAGCGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 2759
Db 2701 TTCGACTCAGATAGCGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 2760
Qy 2760 TTCAGACAGCGACTCAGACAGTGAATTCAGATAGCGACTCAGACTCGGATTCAGCGAGTGA 2819
Db 2761 TTCAGACAGCGACTCAGACAGTGAATTCAGATAGCGACTCAGACTCGGATTCAGCGAGTGA 2820

Db 6447 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAG 6388
Qy 480 TAGCGTTAGTCTGCACCTAAACACAGACGACACAAAACGTGAGTGATACCTAAAACATCGTC 539
Db 6387 TAGGTTAGTCTGCACCTAAACACAGACGACACAAAACGTGAGTGATACCTAAAACATCGTC 6328
Qy 540 AAAACACTAATAATGGCGAAACAGAGTGTGGCGCAAAATCCAGCACAAACAGGAAACGACACA 599
Db 6327 AAACACTAATAATGGCGAAACAGAGTGTGGCGCAAAATCCAGCACAAACAGGAAACGACACA 6268
Qy 600 ATCATCATCAAAATGCAACTACGGAAGAAACGCCGTAACCTGGTGAAGCTACTACTAC 659
Db 6267 ATCATCATCAAAATGCAACTACGGAAGAAACGCCGTAACCTGGTGAAGCTACTACTAC 6208
Qy 660 GACAAAGAAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGGGAGGA 719
Db 6207 GACAAAGAAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGGGAGGA 6148
Qy 720 ATTAGTGAATCAAAACAAAGTAATGAACGACATTTTAATGATACTAATACAGTATCATCTGT 779
Db 6147 ATTAGTGAATCAAAACAAAGTAATGAACGACATTTTAATGATACTAATACAGTATCATCTGT 6088
Qy 780 AAATTCACCTCAAAATTTACAAATGGGAAATGTTTCAACACGCGAAGATCTTCAAC 839
Db 6087 AAATTCACCTCAAAATTTACAAATGGGAAATGTTTCAACACGCGAAGATCTTCAAC 6028
Qy 840 TGAAGCAACACCTTCAAAACAAATGAATCAAGCTCCACAGAGTACAGATGCAAGTAATAAGA 899
Db 6027 TGAAGCAACACCTTCAAAACAAATGAATCAAGCTCCACAGAGTACAGATGCAAGTAATAAGA 5968
Qy 900 TGTAGTGAATCAACGCGTTAATAACAAGTGGCCCTAGAAATGAGACATTTAGTTAGCGGC 959
Db 5967 TGTAGTGAATCAACGCGTTAATAACAAGTGGCCCTAGAAATGAGACATTTAGTTAGCGGC 5908
Qy 960 AGTAGCTGCAGATGACCGGACGCTGGCACAGATATTACGAATCAGTTGACGAATGTGAC 1019
Db 5907 AGTAGCTGCAGATGACCGGACGCTGGCACAGATATTACGAATCAGTTGACGAATGTGAC 5848
Qy 1020 AGTTGGTATTGACCTCTGGTACGATCTGTGTATCCGACCAACAGCAGGTTATGTCAAACTGAA 1079
Db 5847 AGTTGGTATTGACCTCTGGTACGATCTGTGTATCCGACCAACAGCAGGTTATGTCAAACTGAA 5788
Qy 1080 TTATGGTTTTTCACTGCGCTAATCTGTGTGTTTAAAGGTGACACATTCAAAATAAATGTGAC 1139
Db 5787 TTATGGTTTTTCACTGCGCTAATCTGTGTGTTTAAAGGTGACACATTCAAAATAAATGTGAC 5728
Qy 1140 TAAAGAAATTAACCTTAATGGTGTAACTTCAACTGCTTAAAGTGCACCAATTAATGGCTGG 1199
Db 5727 TAAAGAAATTAACCTTAATGGTGTAACTTCAACTGCTTAAAGTGCACCAATTAATGGCTGG 5668
Qy 1200 AGATCAAGTATTGGCAATGGTGAATTCGATAGTGTGATGTTATTTATACATTTAC 1259
Db 5667 AGATCAAGTATTGGCAATGGTGAATTCGATAGTGTGATGTTATTTATACATTTAC 5608
Qy 1260 AGACTATGTAATACTAAAGATGATGTAAGAGAACTTTGACCAATGGCCGCTTATATGA 1319
Db 5607 AGACTATGTAATACTAAAGATGATGTAAGAGAACTTTGACCAATGGCCGCTTATATGA 5548
Qy 1320 CCCTGAAAACTGTTAAAGACAGGTAAATGTCACATTTGCTACTGTCATAGTGTAGTACAAAC 1379
Db 5547 CCCTGAAAACTGTTAAAGACAGGTAAATGTCACATTTGCTACTGTCATAGTGTAGTACAAAC 5488
Qy 1380 AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAAGTTTATATACTTATCTAT 1439
Db 5487 AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAAGTTTATATACTTATCTAT 5428
Qy 1440 TAAAGGTACAAATGACCAAAATCGATATAAACHAATAATACGTTATCGTACGACAAATTTATGT 1499
Db 5427 TAAAGGTACAAATGACCAAAATCGATATAAACHAATAATACGTTATCGTACGACAAATTTATGT 5368
Qy 1500 CAATCCAAAGTGGAGATTAACGTTATTGGCCGGTTTTTAAACAGGTAATTTAAACCAAAATAC 1559
Db 5367 CAATCCAAAGTGGAGATTAACGTTATTGGCCGGTTTTTAAACAGGTAATTTAAACCAAAATAC 5308

Qy 1560 GGATAGTAATGCATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAA 1619
Db 5307 GGATAGTAATGCATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAA 5248
Qy 1620 TGCAGCTGATTTATCTGAAAGTTACTTTGTGAAATCCAGAAACCTTTGAGGATGTCACTAA 1679
Db 5247 TGCAGCTGATTTATCTGAAAGTTACTTTGTGAAATCCAGAAACCTTTGAGGATGTCACTAA 5188
Qy 1680 TAGTGTGAATATTAACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTATACGCTGA 1739
Db 5187 TAGTGTGAATATTAACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTATACGCTGA 5128
Qy 1740 TGAATCAATTAACAACACCGTATATAGTAGTTGTTTAATGGTCAATATGATCCGAATAGCAA 1799
Db 5127 TGAATCAATTAACAACACCGTATATAGTAGTTGTTTAATGGTCAATATGATCCGAATAGCAA 5068
Qy 1800 AGTGATTTAGTGTTCAGTTCAACTTTATATATGGGTATAAATCGAATATAATTTGGCGCTC 1859
Db 5067 AGTGATTTAGTGTTCAGTTCAACTTTATATATGGGTATAAATCGAATATAATTTGGCGCTC 5008
Qy 1860 TATGTCATGGGACAAACGAAGTAGCATTTAATAACGATCAGGTTCTGCTGACGGTATCGA 1919
Db 5007 TATGTCATGGGACAAACGAAGTAGCATTTAATAACGATCAGGTTCTGCTGACGGTATCGA 4948
Qy 1920 TAAACCAAGTTGTTCTCTGAAACCAACCTGATGAGCCCTGGTGAAATTTGAACCAATTCAGAGGA 1979
Db 4947 TAAACCAAGTTGTTCTCTGAAACCAACCTGATGAGCCCTGGTGAAATTTGAACCAATTCAGAGGA 4888
Qy 1980 TTCAAGTTCTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTTCA 2039
Db 4887 TTCAAGTTCTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTTCA 4828
Qy 2040 TTCCGGTAGTGATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGC 2099
Db 4827 TTCCGGTAGTGATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGC 4768
Qy 2100 AAGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGATTTCCGACTCAGCGAGCGA 2159
Db 4767 AAGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGATTTCCGACTCAGCGAGCGA 4708
Qy 2160 TTCCGACTCAGCAAAATGACCTCGGATTCAGATAGCGATTTCTGACTCAGACAGTACTCAGA 2219
Db 4707 TTCCGACTCAGCAAAATGACCTCGGATTCAGATAGCGATTTCTGACTCAGACAGTACTCAGA 4648
Qy 2220 TTCCGACAGTGACTCAGATTCAGATAGCGATTTCTGACTCAGACAGTACTCAGATTCAGA 2279
Db 4647 TTCCGACAGTGACTCAGATTCAGATAGCGATTTCTGACTCAGACAGTACTCAGATTCAGA 4588
Qy 2280 TAGCGATTCAGATTCAGATAGCGATTCAGATTCAGCAAGCGATTTCCGACTCAGACAGCGA 2339
Db 4587 TAGCGATTCAGATTCAGATAGCGATTCAGATTCAGCAAGCGATTTCCGACTCAGACAGCGA 4528
Qy 2340 TTCTGACTCCGACAGTGTATTCGACTCAGACAGCGATTCAGATTCGACAGTGTATTCGA 2399
Db 4527 TTCTGACTCCGACAGTGTATTCGACTCAGACAGCGATTCAGATTCGACAGTGTATTCGA 4468
Qy 2400 CTCAGATAGCGATTTCCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGA 2459
Db 4467 CTCAGATAGCGATTTCCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGA 4408
Qy 2460 CAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGTACTCAGATTCGACAGTGA 2519
Db 4407 CAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGTACTCAGATTCGACAGTGA 4348
Qy 2520 CTCCGATTCAGATAGCGATTCAGATTCGACAGTGTACTCAGATTCGACAGTGTACTCAGA 2579
Db 4347 CTCCGATTCAGATAGCGATTCAGATTCGACAGTGTACTCAGATTCGACAGTGTACTCAGA 4288
Qy 2580 CTCAGACAGTGTATTCGAGTTTCAGCGAGTGATTCGATTCAGATAGTGTATTCGACTCCGA 2639
Db 4287 CTCAGACAGTGTATTCGAGTTTCAGCGAGTGATTCGATTCAGATAGTGTATTCGACTCCGA 4228

Qy	900	TGTAGTTAATCAAGCGGTTAATACAAAGTGCCTAGAAATGAGAGCAATTAGTTAGCGGC	959	Db	3184	TTTCAGATTCTGACCCAGGTTTCAGATTCTGCGACGCGATTCTTAATTTCAGATAGCGGTTTCAGA	3243
Db	2104	TGTAGTTAATCAAGCGGTTAATACAAAGTGCCTAGAAATGAGAGCAATTAGTTAGCGGC	2163	Qy	2040	TTCCGGTAGTGAATTCATCATCAGATAGTGGTTTCAGATTTCAGCGAGTGAATTCAGATTTCAGC	2099
Qy	960	AGTAGCTGCAGATGACCGGACGCTGGGCACAGATATATACGAATCAGTTGACGAATGTGAC	1019	Db	3244	TTCCGGTAGTGAATTCATCATCAGATAGTGGTTTCAGATTTCAGCGAGTGAATTCAGATTTCAGC	3303
Db	2164	AGTAGCTGCAGATGACCGGACGCTGGGCACAGATATATACGAATCAGTTGACGAATGTGAC	2223	Qy	2100	AAAGTGATTTCAGACTTCAGCGAGTGAATTCAGATTTCAGCAAGCGGATTCCGACTTCAGCGAGCGA	2159
Qy	1020	AGTTGGTATTTCAGCTCTGCTAGCAGCTGTCTATCCGACCAAGCAGGTTATGTCAAACTGAA	1079	Db	3304	AAAGTGATTTCAGACTTCAGCGAGTGAATTCAGATTTCAGCAAGCGGATTCCGACTTCAGCGAGCGA	3363
Db	2224	AGTTGGTATTTCAGCTCTGCTAGCAGCTGTCTATCCGACCAAGCAGGTTATGTCAAACTGAA	2283	Qy	2160	TTCCGACTTCAGACAATGACTCCGATTTCAGATAGCGATTTCAGATAGCGATTTCGACTTCAGACAGTGA	2219
Qy	1080	TTATGGTTTTTCAGTGCCCTAAATTCGTCTGTTTAAAGGTGACACATTCAAAAATACTGTACC	1139	Db	3364	TTCCGACTTCAGACAATGACTCCGATTTCAGATAGCGATTTCGACTTCAGACAGTGA	3423
Db	2284	TTATGGTTTTTCAGTGCCCTAAATTCGTCTGTTTAAAGGTGACACATTCAAAAATACTGTACC	2343	Qy	2220	TTCCGACAGTGAATTCAGATTTCAGATAGCGATTTCGACTTCAGACAGTGAATTCAGATTTCAGA	2279
Qy	1140	TAAAGAAATTAACCTTAATGTGTAACTTTCAACTGCTTAAAGTGCACCAATTTATGCAATTTAC	1199	Db	3424	TTCCGACAGTGAATTCAGATTTCAGATAGCGATTTCGACTTCAGACAGTGAATTCAGATTTCAGA	3483
Db	2344	TAAAGAAATTAACCTTAATGTGTAACTTTCAACTGCTTAAAGTGCACCAATTTATGCAATTTAC	2403	Qy	2280	TAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGCAAGTGAATTCGCGACTTCAGACAGCGA	2339
Qy	1200	AGATCAAGTATTGGCAATGCTGTAATCGATAGTGAATGCTTAATGTTATATACATTTAC	1259	Db	3484	TAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGCAAGTGAATTCGCGACTTCAGACAGCGA	3543
Db	2404	AGATCAAGTATTGGCAATGCTGTAATCGATAGTGAATGCTTAATGTTATATACATTTAC	2463	Qy	2340	TTCTGACTTCGACAGTGAATTCGCACTCAGACAGCGATTTCAGATTTCGCAAGTGAATTCGCA	2399
Qy	1260	AGACTATGTAATACTAAAGATGATGTAAAGCAACTTTGACCAATTTGACCAATTTATTTGA	1319	Db	3544	TTCTGACTTCGACAGTGAATTCGCACTCAGACAGCGATTTCAGATTTCGCAAGTGAATTCGCA	3603
Db	2464	AGACTATGTAATACTAAAGATGATGTAAAGCAACTTTGACCAATTTGACCAATTTATTTGA	2523	Qy	2400	CTCAGATAGCGATTTCGCACTCAGATAGCGATTTCAGATTTCAGACAGCGATTTCAGATTTCAGA	2459
Qy	1320	CCCTGAAATGTTTAAAGACAGGTAATGTGACATTGGCTACTGGCATAGTAGTACAAAC	1379	Db	3604	CTCAGATAGCGATTTCGCACTCAGATAGCGATTTCAGATTTCAGACAGCGATTTCAGATTTCAGA	3663
Db	2524	CCCTGAAATGTTTAAAGACAGGTAATGTGACATTGGCTACTGGCATAGTAGTACAAAC	2583	Qy	2460	CAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGCAAGTGAATTCGCGACTTCAGATTTCGCAAGTGA	2519
Qy	1380	AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGCTAAGTTTTTATACTTATCTAT	1439	Db	3664	CAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGCAAGTGAATTCGCGACTTCAGATTTCGCAAGTGA	3723
Db	2584	AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGCTAAGTTTTTATACTTATCTAT	2643	Qy	2520	CTCGGATTTCAGATAGCGATTTCAGATTTCGCAAGTGAATTCAGATTTCGCAAGTGAATTCAGA	2579
Qy	1440	TAAAGGTAACAATTGACCAAAATCGATAAAACAAATTAACGATCGTCAGACAAATTTATGT	1499	Db	3724	CTCGGATTTCAGATAGCGATTTCAGATTTCGCAAGTGAATTCAGATTTCGCAAGTGAATTCAGA	3783
Db	2644	TAAAGGTAACAATTGACCAAAATCGATAAAACAAATTAACGATCGTCAGACAAATTTATGT	2703	Qy	2580	CTCAGACAGTGAATTCGCGATTTCAGCGAGTGAATTCGCGATTTCAGATAGTGAATTCGCGACTCCGA	2639
Qy	1500	CAATCCAGTGGAGATACGTTATTTGCGCCGGTTTTTAACAGGTAATTTAAACCAAAATAC	1559	Db	3784	CTCAGACAGTGAATTCGCGATTTCAGCGAGTGAATTCGCGATTTCAGATAGTGAATTCGCGACTCCGA	3843
Db	2704	CAATCCAGTGGAGATACGTTATTTGCGCCGGTTTTTAACAGGTAATTTAAACCAAAATAC	2763	Qy	2640	CAGTGACTTCGGAATTCAGATAGCGACTCAGACTTCGCGATAGCGACTTCGCGATTTCAGATAGCGA	2699
Qy	1560	GGATAGTAATGCATTAATAGATCAGCAAAATACAAAGTATTAAGTATATAAGTAGATAA	1619	Db	3844	CAGTGACTTCGGAATTCAGATAGCGACTCAGACTTCGCGATAGCGACTTCGCGATTTCAGATAGCGA	3903
Db	2764	GGATAGTAATGCATTAATAGATCAGCAAAATACAAAGTATTAAGTATATAAGTAGATAA	2823	Qy	2700	TTCCGACTTCAGATAGCGATTTCAGAAATCAGACAGCGATTTCAGAAATTCAGACAGCGATTTCAGA	2759
Qy	1620	TGCAGCTGATTATCTGAAAGTTACTTTGTGAATCCAGAAAACTTTGAGGATGTCACTPAA	1679	Db	3904	TTCCGACTTCAGATAGCGATTTCAGAAATCAGACAGCGATTTCAGAAATTCAGACAGCGATTTCAGA	3945
Db	2824	TGCAGCTGATTATCTGAAAGTTACTTTGTGAATCCAGAAAACTTTGAGGATGTCACTPAA	2883	Qy	2760	TTTCAGACAGCGACTTCAGACAGTGAATTCAGATTTCAGATAGTGAATTCGCGATTTCAGCGAGTGA	2819
Qy	1680	TAGTGCATATTTACATTTCCAAATCCAAATCAATATAAGTAGAGTTTAAATAGCCCTGA	1739	Db	3946	TTTCAGACAGCGACTTCAGACAGTGAATTCAGATTTCAGATAGTGAATTCGCGATTTCAGCGAGTGA	4005
Db	2884	TAGTGCATATTTACATTTCCAAATCCAAATCAATATAAGTAGAGTTTAAATAGCCCTGA	2943	Qy	2820	TTTCAGACTTCAGGTTAGTCACTCCGATTTCATCAAGTGAATTCGCGACTTCAGAAAAGTGAATTCAAA	2879
Qy	1740	TGATCAAAATTAACAACCGGTATATAGTAGTTGTTAAAGTGCATATTTGATCCGAATAGCAA	1799	Db	4006	TTTCAGACTTCAGGTTAGTCACTCCGATTTCATCAAGTGAATTCGCGACTTCAGAAAAGTGAATTCAAA	4065
Db	2944	TGATCAAAATTAACAACCGGTATATAGTAGTTGTTAAAGTGCATATTTGATCCGAATAGCAA	3003	Qy	2880	TAGCGATTTCGAGTCAAGGTTCTTAAATATATAGTTCGCGCTTAATTCACCTTAAATATGG	2939
Qy	1800	AGGTGATTTAGCTTTACGTTCAACTTTATATATGGGTATAAATTCGGAATATAATTTGGCGCTC	1859	Db	4066	TAGCGATTTCGAGTCAAGGTTCTTAAATATATAGTTCGCGCTTAATTCACCTTAAATATGG	4125
Db	3004	AGGTGATTTAGCTTTACGTTCAACTTTATATATGGGTATAAATTCGGAATATAATTTGGCGCTC	3063	Qy	2940	TACTTAATGCTTCTAATAAATAAGGCTTAAAGATAGTAAAGAACCAATTCACAGATACAGG	2999
Qy	1860	TATGTCATGGGACCAAGTAGCATTTAATAACGATCAGGTTCTGGTGACGCTATCGA	1919	Db	4126	TACTTAATGCTTCTAATAAATAAGGCTTAAAGATAGTAAAGAACCAATTCACAGATACAGG	4185
Db	3064	TATGTCATGGGACCAAGTAGCATTTAATAACGATCAGGTTCTGGTGACGCTATCGA	3123	Qy	3000	TTCTGAAGATGAACAAATAGCTCACTAAATTTGGGGATTATTAGCATCAATAGGTTTCATT	3059
Qy	1920	TAAACCAAGTTTCTCCTGAAACCACTGATGACGCTGGTGAATTTGAACCAATTTCCAGAGGA	1979	Db	4186	TTCTGAAGATGAACAAATAGCTCACTAAATTTGGGGATTATTAGCATCAATAGGTTTCATT	4245
Db	3124	TAAACCAAGTTTCTCCTGAAACCACTGATGACGCTGGTGAATTTGAACCAATTTCCAGAGGA	3183	Qy	3060	ACTTACTTTTCAGAGAGAAAAAGAAATAAGATAAGAAATAAGTAATAATGATATTAAA	3119
Qy	1980	TTTCAGATTCTGACCCAGGTTTCAGATTCTGGCAGCGATTCTTAATTCAGATAGCGGTTTCAGA	2039				

Db 4246 ACTACTTTTCAGAGAAAAAGAAAAATAAGATAAGAAAAATAAGTAATAATGATATATAA 4305
QY 3120 TTAATCATATGATTCATGAGAGAGCCACCTTAAAGGTCCTTCTTTTACTTGGATTTCC 3179
Db 4306 TTAATCATATGATTCATGAGNA-ACRCCTTAAAGGTCCTTCTTTTACTTGGATTTCC 4364
QY 3180 AAATATATTTTGAATATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 3239
Db 4365 AAATATATTTTGAATATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 4424
QY 3240 GTAGATGTTATATAATTTGGCTTGGCGAAAAATAGGGTGTAGGTAGGTGTTAATTAAG 3299
Db 4425 GTAGATGTTATATAATTTGGCTTGGCGAAAAATAGGGTGTAGGTAGGTGTTAATTAAG 4484
QY 3300 GGAATTAAGAGAAAAATACAGTTGAAAAATAAATTTGCTAGTTTATCATTTGGGAGCAT 3359
Db 4485 GGAATTAAGAGAAAAATACAGTTGAAAAATAAATTTGCTAGTTTATCATTTGGGAGCAT 4544
QY 3360 TATGTGTATCACAATTTTGGGAAAGTAATCGTCGAGTGCAGTGGTTCTGGGGAAGA 3419
Db 4545 TATGTGTATCACAATTTTGGGAAAGTAATCGTCGAGTGCAGTGGTTCTGGGGAAGA 4604
QY 3420 ATCCATATGATCTGAGTGGTGAATGCTGAATTAATAATAATAATAATAATAATAATA 3479
Db 4605 ATCCATATGATCT-AGTGGTTGAATGCTGAATTAATAATAATAATAATAATAATAATA 4663
QY 3480 AAGAGTATAAGAAAGCTT 3498
Db 4664 AAGAGTATAAGAAAGATT 4682

RESULT 6

BX571857_08

WPCOMMENT

Sequence split into 28 fragments LOCUS BX571857 Accession BX571857

Fragment Name	Begin	End
BX571857_00	1	110000
BX571857_01	100001	210000
BX571857_02	200001	310000
BX571857_03	300001	410000
BX571857_04	400001	510000
BX571857_05	500001	610000
BX571857_06	600001	710000
BX571857_07	700001	810000
BX571857_08	800001	910000
BX571857_09	900001	1010000
BX571857_10	10000001	11100000
BX571857_11	11000001	12100000
BX571857_12	12000001	13100000
BX571857_13	13000001	14100000
BX571857_14	14000001	15100000
BX571857_15	15000001	16100000
BX571857_16	16000001	17100000
BX571857_17	17000001	18100000
BX571857_18	18000001	19100000
BX571857_19	19000001	20100000
BX571857_20	20000001	21100000
BX571857_21	21000001	22100000
BX571857_22	22000001	23100000
BX571857_23	23000001	24100000
BX571857_24	24000001	25100000
BX571857_25	25000001	26100000
BX571857_26	26000001	27100000
BX571857_27	27000001	2799802

Continuation (9 of 28) of BX571857 from base 800001 (BX571857 Staphylococcus aureus str

Query Match 93.7%; Score 3277; DB 1; Length 110000;

Best Local Similarity 96.7%; Pred. No. 0;

Matches 3383; Conservative 0; Mismatches 100; Indels 16; Gaps 3;

QY 1 GGTACCATAAATACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATACATTA 60

Db 40748 GGTACCATAAATACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATACATTA 40807

QY	61	GGTAGAGTTTCATTTAAATAATAAAAAATGTTTTCGAATCAAAATCGTAGCTTGTGCTTTGTA	120
Db	40808	GGTAGAGTTTCATTTAAATAATAAAAAATGTTTTCGAATCAAAATCGTAGCTTGTGCTTTGTA	40867
QY	121	ATTCTTAAATAAGCAATAAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA	180
Db	40868	ATTCTTAAATAAGCAATAAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA	40927
QY	181	TATCGATACAAATTAATTCGTATATGCAATTTTAGTGTATATTAATTCATTAACAGAGATT	240
Db	40928	TATCGATACAAATTAATTCGTATATGCAATTTTAGTGTATATTAATTCATTAACAGAGATT	40987
QY	241	AAATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTAAAAAGAGGGGAATAA	299
Db	40988	AAATATATCTTTAAAGGGTATATAGTTAATAATAAATGACTTTTAAAAAGAGGGGAATAA	41047
QY	300	AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTC	359
Db	41048	AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTC	41107
QY	360	AGTGCCTTTAGGTACGTTAATCGGTTTTGGACTTACTCAGCAGTAAAGACGATGCAAG	419
Db	41108	AGTGCCTTTAGGTACGTTAATCGGTTTTGGACTTACTCAGCAGTAAAGACGATGCAAG	41167
QY	420	TGAAAAATAGTGTTCGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAAG	479
Db	41168	TGAAAAATAGTGTTCGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAAG	41227
QY	480	TAGCGTTAGTGTGCACTTAAACAGACGACACAAACGTGAGTGATTAATAACATCGTC	539
Db	41228	TAGCGTTAGTGTGCACTTAAACAGACGACACAAACGTGAGTGATTAATAACATCGTC	41287
QY	540	AAACACTAATATGGCGGAAACGAGTGTGGCGGMAAATCCAGCACAACAGAAACGACACA	599
Db	41288	AAACACTAATATGGCGGAAACGAGTGTGGCGGMAAATCCAGCACAACAGAAACGACACA	41347
QY	600	ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC	659
Db	41348	ATCAGCATTAAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAG--CTACTAC	41404
QY	660	GACACGAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATCAAAATCGGAGGA	719
Db	41405	GACACGAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATCAAAATCGGAGGA	41464
QY	720	ATTAGTGAATCAACAAAGTAATGAACGACTTTTAATGATCTACTAATACAGTATCATCTGT	779
Db	41465	ATTAGTGAATCAACAAAGTAATGAACGACTTTTAATGATCTACTAATACAGTATCATCTGT	41524
QY	780	AAATTCACCTCAAAATTTCTACAAATCGGAAATGTTTCAACAAACGCAAGATACTTCAAC	839
Db	41525	AAATTCACCTCAAAATTTCTACAAATCGGAAATGTTTCAACAAACGCAAGATACTTCAAC	41584
QY	840	TGAAGCAACACCTTCAAAACAAATGAATCAGTCCACAGAGTACAGATGCAAGTAATAAGA	899
Db	41585	TGAAGCAACACCTTCAAAACAAATGAATCAGTCCACAGAGTACAGATGCAAGTAATAAGA	41644
QY	900	TGTAGTTAATCAAGCGGTTAATACAGTGGCGCTAGAATGAGAGCAATTTAGTTTACGGC	959
Db	41645	TGTAGTTAATCAAGCGGTTAATACAGTGGCGCTAGAATGAGAGCAATTTAGTTTACGGC	41704
QY	960	AGTAGCTGCAGATGCAACCGGAGCTGGCAGAGATATTACGAATCAGTTGACGAATGTGAC	1019
Db	41705	AGTAGCTGCAGATGCAACCGGAGCTGGCAGAGATATTACGAATCAGTTGACGAATGTGAC	41764
QY	1020	AGTTGGTATTGACTCTGGTACCACTGTGTATCGCACCAAGCAGGTTATGTCAAACTGAA	1079
Db	41765	AGTTGGTATTGACTCTGGGAGATACAGTTTATCGCACCAAGCAGGCTATGTCAAACTGAA	41824
QY	1080	TTATGGTTTTTTCAGTGCCTTAATCTGCTGTTAAAGGTGACACATTTCAAAATACCTGACC	1139
Db	41825	TTATGGTTTTTTCAGTGCCTTAATCTGCTGTTAAAGGTGACACATTTCAAAATACCTGACC	41884

CDS	complement(5619..5786) /gene="MW0534" /note="ORFID:MW0534" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB94399.1" /db_xref="GI:21203700" /translation="MIYRQVHEGAPVVEIITKFFQHSVKCDSDSFDTFIFKLKLSLQDDIDHMKVS" complement(6293..7123) /gene="MW0535" complement(6293..7123) /gene="MW0535" /note="ORFID:MW0535" hypothetical protein, similar to phosphomethylpyrimidine kinase, thid homologue" /codon_start=1 /transl_table=11 /protein_id="BAB94400.1" /db_xref="GI:21203701" /translation="MALKKVLTIAGSDTSAGAGMADLKTFOELDTYGMVALTAVTMKDQWSDHVTPLPMDVPEKOLETALSIGPDALKTGMGTSEIIRKAGEVYASNAQYFVYDPMVCKGEDEVLPNGNTEAMIKYLLPKAIIVTPNLFAGQSLGKUNSIEMKKAAATLIFDKGAOHVLIKGGKALDQDKSYLDYDGGTFYQLTDTDMFQQSYNHGAGCTFAAATTAYLANGSPKPEAVISAKAFVASAIKNGWKMDVFGVDHGAYNRIEHIIDVEVTEV" 7307..7963 /gene="ung" 7307..7963 /gene="ung" /note="ORFID:MW0536" /codon_start=1 /transl_table=11 /product="uracil-DNA glycosylase" /protein_id="BAB94401.1" /db_xref="GI:21203702" /translation="MNEWSQIFPHDITTKHDFKAMHDFLEKEYSTAIVPDRNIYQAFDLTPFNKVLIGQDPDHPGNGHGLAFVOPNAKFPPLSNMNYKELADDIGCVRQTPHLQWAEQVILLNTVITVROGEANSHRDIGWETFDIIEKAVSDYKEHVVFILWKEPAQOKIKLIDTSKHCIKSVHPSPLSAVRGFGSPKSKANNAYLBSVGKSPINWCESEA" 7964..8344 /gene="MW0537" 7964..8344 /gene="MW0537" /note="ORFID:MW0537" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB94402.1" /db_xref="GI:21203703" /translation="MLNRETLIARIEOELVQAEAOHDHDFEKHVAIHILTSIYASTSNTPHLIGEOQNRRIANHNOMPSQITQTPHOVTAABIEANGKVNTHSAHHNKSYSQPSNQORLADDDIDNGESIFDF" 8477..8845 /gene="MW0538" 8477..8845 /gene="MW0538" /note="ORFID:MW0538" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="BAB94403.1" /db_xref="GI:21203704" /translation="MKLFIILGALNNMAYVTGAFGAHGLQKLSIDHYLSWSEKATYYQMHLFIQWMLIITATFRFAG" 8966..10450 /gene="MW0539" 8966..10450	Best Local Similarity 95.8%; Pred. No. 0; Matches 3403; Conservative 0; Mismatches 80; Indels 70; Gaps 4;
gene	1 GGTACCATAAATTACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATTACATTA 60 241201 GGTACCATAAATTACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATTACATTA 241260	
CDS	61 GGTAGAGTTTCATATTAATAAATAAATGTTTGCATCAAAATCGTACGTTGTCGTTGTA 120 241261 GGTAGAGTTTCATATTAATAAATAAATGTTTGCATCAAAATCGTACGTTGTCGTTGTA 241320 121 ATTCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTTGGATAATAAAA 180 241321 ATTTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTTGGATAATAAAA 241380 181 TATCGATACAAATTAATTTGCTATTAATGCAATTTTAGTGTATAATTTCCATTAACAGAGATT 240 241381 TATCGATACAAATTAATTTGCTATTAATGCAATTTTAGTGTATAATTTCCATTAACAGAGATT 241440 241 AAATATATC-TTAAAGGCTATATAGTTAATAATAAATGACTTTTTTAAAGAGGGAATAA 299 241441 AAATATATCTTTAAAGGCTATATAGTTAATAATAAATGACTTTTTTAAAGAGGGAATAA 241500 300 AATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGCCTTC 359 241501 AATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGCCTTC 241560 360 AGTCCTTGTAGGTACGTTAATCGGTTTGGACTACTCAGCAGTAGTAAGAGCAGATGCAAG 419 241561 AGTCCTTGTAGGTACGTTAATCGGTTTGGACTACTCAGCAGTAGTAAGAGCAGATGCAAG 241620 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTATGATTCAG 479 241621 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTATGATTCAG 241680 480 TAGCGTTAGTGTCTGCACCTTAAACAGACGACACAAACGTCGTAGTGATGATCTAAACATCGTC 539 241681 TAGCGTTAGTGTCTGCACCTTAAACAGACGACACAAACGTCGTAGTGATGATCTAAACATCGTC 241740 540 AAACACTAATAATGGCGAAGCAGTGTGGCGCAAAATCCAGCACAACAGGAAACGACACA 599 241741 AAACACTAATAATGGCGAAGCAGTGTGGCGCAAAATCCAGCACAACAGGAAACGACACA 241800 600 ATCATCATCAACAAATGCACTACGGAAGAAACCGCGTAACTCGTGAAGCTACTACTAC 659 241801 ATCAGCATTAACAAATGCACTACGGAAGAAACCTCGGTAACTGGTGAAG---CTACTAC 241857 660 GACAAACGAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGA 719 241858 GGCAAACGAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGA 241917 720 ATTAGTGAATCAACAAAGTAAAGACGACTTTTAATGATACATAATCAGTATCATCTGT 779 241918 ATTAGTGAATCAACAAAGTAAAGACGACTTTTAATGATACATAATCAGTATCATCTGT 241977 780 AAATTCACCTCAAAATTTCTACAAATCGGAAAATGTTTCAACACGCAAGATCTTCAAC 839 241978 AAATTCACCTCAAAATTTCTACAAATCGGAAAATGTTTCAACACGCAAGATCTTCAAC 242037 840 TGAAGCAACACCTTCAAAACAAATGAATCAGTCTCCACAGAGTACAGATCAAGTAATAAGA 899 242038 TGAAGCAACACCTTCAAAACAAATGAATCAGTCTCCACAGAGTACAGATCAAGTAATAAGA 242097 900 TGTAGTTAATCAGCGGTTAATACAGTGGCGCTAGCAATCAGAGCATTTAGTTTAGCGGC 959 242098 TGTAGTTAATCAGCGGTTAATACAGTGGCGCTAGCAATCAGAGCATTTAGTTTAGCGGC 242157 960 AGTAGCTGCAGATGCACCGGCAGCTGGCAGACAGATATTACGAATCAGTTGACGAATGTGAC 1019 242158 AGTAGCTGCAGATGCACCGGCAGCTGGCAGACAGATATTACGAATCAGTTGACGAATGTGAC 242217 1020 AGTTGGTATTGACTCTGGGTACGACTGTGTATCCGCAACAGCAGGTTATGTCAAACTGAA 1079	
gene	7964..8344 /gene="MW0537" 7964..8344 /gene="MW0537" /note="ORFID:MW0537" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB94401.1" /db_xref="GI:21203702" /translation="MNEWSQIFPHDITTKHDFKAMHDFLEKEYSTAIVPDRNIYQAFDLTPFNKVLIGQDPDHPGNGHGLAFVOPNAKFPPLSNMNYKELADDIGCVRQTPHLQWAEQVILLNTVITVROGEANSHRDIGWETFDIIEKAVSDYKEHVVFILWKEPAQOKIKLIDTSKHCIKSVHPSPLSAVRGFGSPKSKANNAYLBSVGKSPINWCESEA" 7964..8344 /gene="MW0537" 7964..8344 /gene="MW0537" /note="ORFID:MW0537" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB94401.1" /db_xref="GI:21203702" /translation="MNEWSQIFPHDITTKHDFKAMHDFLEKEYSTAIVPDRNIYQAFDLTPFNKVLIGQDPDHPGNGHGLAFVOPNAKFPPLSNMNYKELADDIGCVRQTPHLQWAEQVILLNTVITVROGEANSHRDIGWETFDIIEKAVSDYKEHVVFILWKEPAQOKIKLIDTSKHCIKSVHPSPLSAVRGFGSPKSKANNAYLBSVGKSPINWCESEA" 7964..8344 /gene="MW0537" 7964..8344 /gene="MW0537" /note="ORFID:MW0537" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB94401.1" /db_xref="GI:21203702" /translation="MNEWSQIFPHDITTKHDFKAMHDFLEKEYSTAIVPDRNIYQAFDLTPFNKVLIGQDPDHPGNGHGLAFVOPNAKFPPLSNMNYKELADDIGCVRQTPHLQWAEQVILLNTVITVROGEANSHRDIGWETFDIIEKAVSDYKEHVVFILWKEPAQOKIKLIDTSKHCIKSVHPSPLSAVRGFGSPKSKANNAYLBSVGKSPINWCESEA" 7964..8344 /gene="MW0537" 7964..8344 /gene="MW0537" /note="ORFID:MW0537" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB94401.1" /db_xref="GI:21203702" /translation="MNEWSQIFPHDITTKHDFKAMHDFLEKEYSTAIVPDRNIYQAFDLTPFNKVLIGQDPDHPGNGHGLAFVOPNAKFPPLSNMNYKELADDIGCVRQTPHLQWAEQVILLNTVITVROGEANSHRDIGWETFDIIEKAVSDYKEHVVFILWKEPAQOKIKLIDTSKHCIKSVHPSPLSAVRGFGSPKSKANNAYLBSVGKSPINWCESEA"	
CDS	7964..8344 /gene="MW0537" 7964..8344 /gene="MW0537" /note="ORFID:MW0537" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB94401.1" /db_xref="GI:21203702" /translation="MNEWSQIFPHDITTKHDFKAMHDFLEKEYSTAIVPDRNIYQAFDLTPFNKVLIGQDPDHPGNGHGLAFVOPNAKFPPLSNMNYKELADDIGCVRQTPHLQWAEQVILLNTVITVROGEANSHRDIGWETFDIIEKAVSDYKEHVVFILWKEPAQOKIKLIDTSKHCIKSVHPSPLSAVRGFGSPKSKANNAYLBSVGKSPINWCESEA"	
gene	8477..8845 /gene="MW0538" 8477..8845 /gene="MW0538" /note="ORFID:MW0538" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="BAB94402.1" /db_xref="GI:21203703" /translation="MLNRETLIARIEOELVQAEAOHDHDFEKHVAIHILTSIYASTSNTPHLIGEOQNRRIANHNOMPSQITQTPHOVTAABIEANGKVNTHSAHHNKSYSQPSNQORLADDDIDNGESIFDF" 8477..8845 /gene="MW0538" 8477..8845 /gene="MW0538" /note="ORFID:MW0538" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="BAB94402.1" /db_xref="GI:21203703" /translation="MLNRETLIARIEOELVQAEAOHDHDFEKHVAIHILTSIYASTSNTPHLIGEOQNRRIANHNOMPSQITQTPHOVTAABIEANGKVNTHSAHHNKSYSQPSNQORLADDDIDNGESIFDF"	
CDS	8477..8845 /gene="MW0538" 8477..8845 /gene="MW0538" /note="ORFID:MW0538" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="BAB94402.1" /db_xref="GI:21203703" /translation="MLNRETLIARIEOELVQAEAOHDHDFEKHVAIHILTSIYASTSNTPHLIGEOQNRRIANHNOMPSQITQTPHOVTAABIEANGKVNTHSAHHNKSYSQPSNQORLADDDIDNGESIFDF"	
gene	8477..8845 /gene="MW0538" 8477..8845 /gene="MW0538" /note="ORFID:MW0538" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="BAB94402.1" /db_xref="GI:21203703" /translation="MLNRETLIARIEOELVQAEAOHDHDFEKHVAIHILTSIYASTSNTPHLIGEOQNRRIANHNOMPSQITQTPHOVTAABIEANGKVNTHSAHHNKSYSQPSNQORLADDDIDNGESIFDF"	
CDS	8477..8845 /gene="MW0538" 8477..8845 /gene="MW0538" /note="ORFID:MW0538" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="BAB94402.1" /db_xref="GI:21203703" /translation="MLNRETLIARIEOELVQAEAOHDHDFEKHVAIHILTSIYASTSNTPHLIGEOQNRRIANHNOMPSQITQTPHOVTAABIEANGKVNTHSAHHNKSYSQPSNQORLADDDIDNGESIFDF"	
gene	8966..10450 /gene="MW0539" 8966..10450	
CDS	8966..10450	
Query Match	92.8%; Score 3245; DB 1; Length 290150;	

Db	242218	AGTTGGTATTGACTCTGGAGATACAGTCTTATCCGACCAAGCAGCGCTATGTCAAACCTGAA	242277
Qy	1080	TTATGGTTTTTTCAGTGCCTAAATCTCTGCTTTAAAGGTGACACATTCAAAATAAAGTGTACC	1139
Db	242278	TTATGGTTTTTTCAGTGCCTAAATCTCTGCTTTAAAGGTGACACATTCAAAATAAAGTGTACC	242337
Qy	1140	TAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTGCACCAATTAATGCTGG	1199
Db	242338	TAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTGCCTCAAATTAATGSCCGG	242397
Qy	1200	AGATCAAGTATTGGCAATGGTGTAAATCGATAGTGTAAATGTTTATTTATACATTTAC	1259
Db	242398	AGATCAAGTATTGGCAATGGTGTAAATCGATAGTGTAAATGTTTATTTATACATTTAC	242457
Qy	1260	AGACTATGTAAATPACTAAAGATGATGTAAAGCAACTTTGACCATGCCGCTTATATTGA	1319
Db	242458	AGACTATGTGTACTAAAGAAAATGTAAACAGCTAATATTACTATGCCAGCTTATATTGA	242517
Qy	1320	CCCTGAAAATGTTAAAGACACAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAAC	1379
Db	242518	CCCTGAAAATGTTTACAAAGACAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAAC	242577
Qy	1380	AGCAAAACAAAACAGTATTAGTAGATTATGAAAAATATGGTAAGTTTTTATAACTTTATCTAT	1439
Db	242578	AGCAAAACAAAACAGTATTAGTAGATTATGAAAAATATGGTAAGTTTTTATAACTTTATCTAT	242637
Qy	1440	TAAAGGTACAAATTGACCAATTCGATAAACAAATAATACGTATCTGTCAGACAATTTATGT	1499
Db	242638	TAAAGGTACAAATTGACCAATTCGATAAACAAATAATACGTATCTGTCAGACAATTTATGT	242697
Qy	1500	CAATCCAGTGGAGATACAGTTATCGCCGGTTTTTAAACAGGTAATTTAAACCAAAATAC	1559
Db	242698	CAATCCAGTGGAGATACAGTTATCGCCGGTTTTTAAACAGGTAATTTAAACCAAAATAC	242757
Qy	1560	GGATAGTAATGCATTATAGATCAGCAAAATACAAAGTATTAAGTATATAAAGTAGATAA	1619
Db	242758	GGATAGTAATGCATTATAGATCAGCAAAATACAAAGTATTAAGTATATAAAGTAGATAA	242817
Qy	1620	TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCACTAA	1679
Db	242818	TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCACTAA	242877
Qy	1680	TAGTGTCAATATTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATAGCCCTGA	1739
Db	242878	TAGTGTCAATATTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATAGCCCTGA	242937
Qy	1740	TGATCAAAATTACAACACCGTATATAGTAGTTGTTTAATGGTCATATTGATCCGAATAGCAA	1799
Db	242938	TGATCAAAATTACAACACCGTATATAGTAGTTGTTTAATGGTCATATTGATCCGAATAGCAA	242997
Qy	1800	AGGTGATTTAGCTTTAGCTTCAACTTTATATGGGTATAACTCGAATATAATTTGGCGCTC	1859
Db	242998	AGGTGATTTAGCTTTAGCTTCAACTTTATATGGGTATAACTCGAATATAATTTGGCGCTC	243057
Qy	1860	TATGTATGGGACACACAGTAGCATTTAAACGGATCAGGTTCTGTGACGGTATCGA	1919
Db	243058	TATGTATGGGACACACAGTAGCATTTAAACGGATCAGGTTCTGTGACGGTATCGA	243117
Qy	1920	TAAACCCAGTTGTTCTGTAACCAACCTGATGAGCCTGGTGAATTTGAACCAATTTCCAGAGGA	1979
Db	243118	TAAACCCAGTTGTTCTGTAACCAACCTGATGAGCCTGGTGAATTTGAACCAATTTCCAGAGGA	243177
Qy	1980	TTTCAGATTTCAACCCAGGTTTCAAGATTTGCGGACGGATTCTTAATTCAGATAGCGGTTTCA	2039
Db	243178	TTTCAGATTTCAACCCAGGTTTCAAGATTTGCGGACGGATTCTTAATTCAGATAGCGGTTTCA	243237
Qy	2040	TTCCGGTGTAGTCTATACATAGATAGTGGTTTCAGATTCAGCGAGTGTTCAGATTCAGC	2099
Db	243238	TTCCGGTGTAGTCTATACATAGATAGTGGTTTCAGATTCAGCGAGTGTTCAGATTCAGC	243297
Qy	2100	AAGTGATTCAGACTCAGCGAGTGTTCAGATTCAGCAAGCGATTCGCGACTCAGCGAGCGA	2159
Db	243298	AAGTGATTCAGACTCAGCGAGTGTTCAGATTCAGCAAGCGATTCGCGACTCAGCGAGCGA	243357
Qy	2160	TTCCGACTCAGCAATGACTCCGATTCAGATAGCAGATTTCTGACTCAGACAGTGAATTCAGA	2219
Db	243358	TTTCAGATTCAGCGAGCGGATTCAGATTCAGCGAGTGTTCAGATTCAGCGAGTGAATTCAGA	243417
Qy	2220	T-----TCCGA	2225
Db	243418	TTTCAGCAAGCGATTCGCACTCAGACAAATGACTCCGATTCAGATAGCAGATTTCTGACTCAGA	243477
Qy	2226	CAGTGAATTCAGATTCAGATAGCGATTCGACTCAGACAGTGAATTCAGATTCAGATAGCGA	2285
Db	243478	CAGTGAATTCAGATTCGATAGCGATTCGACTCAGACAGTGAATTCAGATTCAGATAGCGA	243537
Qy	2286	TTTCAGATTCAGATAGCGATTCGATAGCGATTCGACTCAGACAGTGAATTCGACTCAGACAGGATTTCTGA	2345
Db	243538	TTTCAGATTCAGATAGCGATTCGATAGCGATTCGACTCAGACAGTGAATTCGACTCAGACAGGATTTCTGA	243597
Qy	2346	CTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGCGACAGTGAATTCGACTCAGAC	2405
Db	243598	CTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGCGACAGTGAATTCGACTCAGAC	243657
Qy	2406	TAGCGATTCGCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGA	2465
Db	243658	TAGCGATTCGCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGA	243717
Qy	2466	TTTCAGATTCAGATAGCGATTCGATTCGCGACAGTGAATTCGCGACAGTGAATTCGACTCAGA	2525
Db	243718	TTTCAGATTCAGATAGCGATTCGATTCGCGACAGTGAATTCGCGACAGTGAATTCGACTCAGA	243777
Qy	2526	TTTCAGATAGCGATTCGATTCGCGACAGTGAATTCGCGACAGTGAATTCGCGACAGTGAATTCG	2585
Db	243778	TTTCAGATAGCGATTCGATTCGCGACAGTGAATTCGCGACAGTGAATTCGCGACAGTGAATTCG	243837
Qy	2586	CAGTGAATTCGCGATTCAGCGAGTGAATTCGCGATTCAGATAGTGAATTCGCGACTCCGACAGTGA	2645
Db	243838	CAGTGAATTCGCGATTCAGCGAGTGAATTCGCGATTCAGATAGTGAATTCGCGACTCCGACAGTGA	243897
Qy	2646	CTCCGATTCAGATAGCGACTCAGACTCCGATAGCGACTCCGATTCAGATAGCGACTCCGATTCGGA	2705
Db	243898	CTCCGATTCAGATAGCGACTCAGACTCCGATAGCGACTCCGATTCAGATAGCGACTCCGATTCGGA	243957
Qy	2706	CTTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGATTCAGA	2765
Db	243958	CTTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGATTCAGA	244017
Qy	2766	CAGCGACTCAGACAGTGAATTCAGATTCAGATAGTGAATTCGCGACTCCGCGAGTGAATTCAGA	2825
Db	244018	CAGC-----GACTCAGATTCAGATAGTGAATTCGCGACTCCGCGAGTGAATTCAGA	244065
Qy	2826	CTCAGGTAGTGAATTCAGATTCAGATAGTGAATTCGCGACTCCGCGAGTGAATTCAGATTCAGA	2885
Db	244066	CTCAGGTAGTGAATTCAGATTCAGATAGTGAATTCGCGACTCCGCGAGTGAATTCAGATTCAGA	244125
Qy	2886	TTCCGAGTCAAGTTCTAACCAATTAATAGTGAATTCAGATTCAGATAGTGAATTCAGATTCAGA	2945
Db	244126	TTCCGAGTCAAGTTCTAACCAATTAATAGTGAATTCAGATTCAGATAGTGAATTCAGATTCAGA	244185
Qy	2946	TGCTTCTTAATAAATCAGCGTGAATAGTGAATTAAGAACCTTACAGATTCAGATTCAGATTCAGA	3005
Db	244186	TGCTTCTTAATAAATCAGCGTGAATAGTGAATTAAGAACCTTACAGATTCAGATTCAGATTCAGA	244245
Qy	3006	AGATGAAGCAAAATACGTCACCTAATTTGGGATTTATAGCATCAATAGGTTTCAATTTACTACT	3065
Db	244246	AGATGAAGCAAAATACGTCACCTAATTTGGGATTTATAGCATCAATAGGTTTCAATTTACTACT	244305
Qy	3066	TTTCAGAGCAAAATAGCAAAATAGCAAAATAGCAAAATAGCAAAATAGCAAAATAGCAAAATAGCA	3125
Db	244306	TTTCAGAGCAAAATAGCAAAATAGCAAAATAGCAAAATAGCAAAATAGCAAAATAGCAAAATAGCA	244365
Qy	3126	ATATGATTCATGAAGAGCCACCTTAAAGGTGCTTCTTTTACTTGGATTTTCCAAATAT	3185
Db	244366	ATATGATTCATGAAGAGCCACCTTAAAGGTGCTTCTTTTACTTGGATTTTCCAAATAT	244425


```
/product="MnhD homologue"
/protein_id="BAB56787.1"
/db_xref="GI:14246394"
/translation="MLSNLLILPMLPFLCALILVFLKNDRIISKYLYLGTWTTITII
SLMLIYVQRHPTITLDFGSGAPGQIFGLDSLSLIMVTTASVFITILIMAYGGRGE
HKANRHYLPSPFILSVGVIGSFLTSDLPNLYVMPFIMLLASFLVITLQSQVEOLRAA
IIVVNLIGSWFLGLGLGLYKTVGTILNFHSHARLNDMGDNRTVMISLIFLVAFS
AKAALVPMFLPKAVANLTAEALFAALMTKVGAYALIRFPTLLFDORNLHPELLA
TMAATIMVIGIAYDKIKKIAAYQVILSIGFIIILGLGNTFAGINGALFYLVNDI
VKTLLFIIGSLVITGVYQYVLYNGLAKEPFLGVAFIIMFAIGVPPFSGPPGK
VLIFGALQNGNYIGLAMIITSLIAMSFLRPFYFMYFGDKGSEVNFKKIPLYRKR
ILSLVVVIAIGIAAPVVLNVTSDEALNTSDQLYQKLVNPHLKGED"
5327..5809
/locus_tag="SAV0626"
5327..5809
/locus_tag="SAV0626"
/codon_start=1
/transl_table=11
/product="Na+ antiporter"
/protein_id="BAB56788.1"
/db_xref="GI:14246395"
/translation="WNQIVLNIIIFLWVLFQDEDFKPFSTFFSGYLIGLIVYILHR
FSDDFYVRKIWAITKFLGVLYQLITSSITINVLFKDKMNFGLLSYETRLTSDW
AITEFLIILITPGSVIRISQDSKKFTHSIDVSEKEKDSLRSIKHYEDLILEVSR
"
5806..6108
/locus_tag="SAV0627"
5806..6108
/locus_tag="SAV0627"
/codon_start=1
/transl_table=11
/product="similar to Na+ antiporter"
/protein_id="BAB56789.1"
/db_xref="GI:14246396"
/translation="MIQTTHIMLIISLIIFGIALIICLPRIKGPFTADRVVTFDIT
SAVMSIVGLSVLMTGVFLDSIMLIAISFVSSVSIREFIGGHVFNNGNKNRL"
6083..6520
/locus_tag="SAV0628"
6083..6520
/locus_tag="SAV0628"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB56790.1"
/db_xref="GI:14246397"
/translation="MBITKEIFSLIAAVMLLLGSFIALISAIGIVKFDVFLRSHAAT
KSTLSVLTLIGLVIFVNTGTFESVRLSLVFINLTPVGMHLVARAAYRNGAYM
YRKNDATHASILLSSNEQSTALQLRAKKREHRKKWYQND"
6860..8902
/locus_tag="SAV0629"
6860..8902
/locus_tag="SAV0629"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB56791.1"
/db_xref="GI:14246398"
/translation="MEIFETPILIFIAVVLISFVHTTIPKVPFLAFIQIFLGMFLFTTP
IPVQNFDSFLFMTMIAPIFLFVGVNSVRVHLRKIPKVMMLGLVITTVIGVGLF
IHWIWPDLPGAFAIAILCPTDAVAQAITKGKLPKGMAMTILEGESLINDAAGII
SFKIAGVLTGTASLDVAVQLFIALSIGGAVVGLLIGMALVRFRLTMRGYNINM
FTIILQTPFTFYLIALFELFHASGIIAAVAGLVGHGFERDMQVRLQMSYHFWNI
LGYVLTGIFVILGVPFVFIKTEPHNLIFLIGITIVLAVLFRFVWYVLY
PVFYLAIIPFOKMTKDDNDPTTKPKPRSLVALIMTLGCVHGTISIALATLTPFL
AGHAFYRNDLLFIASGMVILISLVAQVLLPLTKPKPTVIGNMSPKVARIVILEQ
VIDINKSTFETSFKYGNVIEKHDKLAFILKTEKODENSEKELERQKIAFNVEKTI
LESIVDEQIITNSVLENTYRABRTQVTRQASLIRMLVLRGALLKRVQTRVNSAS
SLSVTDNLMEINKLKVHYNVSRSLSKETIKDNTLEIGVMVCDGVMRIENLTSPNFF
NSASEDTTKIKLNAIRQRILRELIDTDEVSEGTALKLREALINDEMIIVDSMT"
complement(9010..10329)
/gene="tnp"
/locus_tag="SAV0630"
complement(9010..10329)
```

```
/gene="tnp"
/locus_tag="SAV0630"
/codon_start=1
/transl_table=11
/product="transposase for IS1181"
/protein_id="BAB56792.1"
/db_xref="GI:14246399"
/translation="WCNDTLELLRIKDENIKYINQIEDVILKGGKATVVMNAVLYTKPS
"
Query Match 84.6%; Score 2960.6; DB 1; Length 348477;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 3179; Conservative 0; Mismatches 319; Indels 7; Gaps 2;
QY 1 GGTACATAAATTACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTCATTA 60
DB 193322 GGTACATAAATTACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTCATTA 193381
QY 61 GGTAGAGTTTCATTAATAATAAAAAATGTTTGCAATCAATCGTACGTTGCGTTGTA 120
DB 193382 GGTAGAGTTTCATTAATAATAAAAAATGTTTGCAATCAATCGTACGTTGCGTTGTA 193441
QY 121 ATTCTTAATAGCAATAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA 180
DB 193442 ATTCTTAATAGCAATAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA 193501
QY 181 TATCGATACAAATTAATTCCTATAATGCAATTTTAGTGTATATATCCATTAAACAGAGTT 240
DB 193502 TATCGATACAAATTAATTCCTATAATGCAATTTTAGTGTATATATCCATTAAACAGAGTT 193561
QY 241 AAATATATTC-TTAAAGGGTATATAGTTAATAATAAATGATTTTAAAAAGAGGGATTA 299
DB 193562 AAATAATCTTTAAAGGGTATATAGTTAATAATAAATGATTTTAAAAAGAGGGATTA 193621
QY 300 AATGAATATGAAGAAAAGAAAACACGCAATTCGAAAAAATTCGATTTGGCGTGGCTTC 359
DB 193622 AATGAATATGAAGAAAAGAAAACACGCAATTCGAAAAAATTCGATTTGGCGTGGCTTC 193681
QY 360 AGTGCTTTGTAGGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAAGAGCAGATGCAAG 419
DB 193682 AGTGCTTTGTAGGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAAGAGCAGATGCAAG 193741
QY 420 TGAATAATAGTGTTCGCAATCTGATAGCGCAAGTACGAAAGCAAAAGTAATGATTCAG 479
DB 193742 TGAATAATAGTGTTCGCAATCTGATAGCGCAAGTACGAAAGCAAAAGTAATGATTCAG 193801
QY 480 TAGCGTTAGTGTGTCACCTTAACACAGCAGACAAACGTCGAGTGATATAAAACATCGTC 539
DB 193802 TAGCGTTAGTGTGTCACCTTAACACAGCAGACAAACGTCGAGTGATATAAAACATCGTC 193861
QY 540 AAACACATTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCACAACAGGAAACGACACA 599
DB 193862 AAACACATTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCACAACAGGAAACGACACA 193921
QY 600 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGCTGAAGTACTACTAC 659
DB 193922 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGCTGAAGTACTACTAC 193981
QY 660 GACAAAGAAATCAAGCTAATACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGGA 719
DB 193982 GACAAAGAAATCAAGCTAATACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGGA 194041
QY 720 ATTAGTGAATCAAAACAAGTAATGAAACGACTTTTAAATGATACTAATACAGTATCATCTGT 779
DB 194042 ATTAGTGAATCAAAACAAGTAATGAAACGACTTTTAAATGATACTAATACAGTATCATCTGT 194101
QY 780 AAATTCACCTCAAAATTTCAAAATCGGAAAAATGTTTCAACACGCAAGTACTTCAAC 839
DB 194102 AAATTCACCTCAAAATTTCAAAATCGGAAAAATGTTTCAACACGCAAGTACTTCAAC 194161
QY 840 TGAAGCAACACCTTCAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTATAAAGA 899
DB 194162 TGAAGCAACACCTTCAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTATAAAGA 194221
```

QY	900	TGTAGTTAATCAAGCGTTAAATACAAAGTGGCTAGAAATGACGAGCAATTTAGTTACGGC	959	Db	195302	TTCCAGATTCGACCCAGGTTCCAGATTCGCGCGGATTTCTAATTCAGATAGCGGTTCCAGA	195361
Db	194222	TGTAGTTAGTCAAGCGGTTAAATCCAAAGTACGCGCTAGAAATGAGACAAATTTAGTTACGGC	194281	QY	2040	TTCCGGTAGTGAATTCACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGC	2099
QY	960	AGTAGTCGAGATGCAACCGGAGCTGGCACAGATATTACGAATTCAGATTCAGCGAGTGAATTCAGATTCAGC	1019	Db	195362	TTTCGCGAGTGAATTCACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGC	195421
Db	194282	AGTAGTCGAGATGCAACCGGAGCTGGCACAGATATTACGAATTCAGATTCAGCGAGTGAATTCAGATTCAGC	194341	QY	2100	AAAGTGAATTCAGACTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGATTCAGCGAGGCA	2159
QY	1020	AGTTGGTATGTACTCTGGTACGACTGTGTATCGCACCAAGCAGGTTATGTCAAACTGAA	1079	Db	195422	AAAGTGAATTCAGACTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGATTCAGCGAGGCA	195481
Db	194342	AGTTGGTATGTACTCTGGTACGACTGTGTATCGCACCAAGCAGGTTATGTCAAACTGAA	194401	QY	2160	TTCCGAGTTCAGAAATGACTCGGATTCAGATAGCGATTCGACTCAGACAGTGAATTCAGATTCAGC	2219
QY	1080	TTATGCTTTTTCAGTGGCTAAATCTGCTGTTAAAGGTGACATTCACAAATTAACCTGAC	1139	Db	195482	TTCCGAGTTCAGAAATGACTCGGATTCAGATAGCGATTCGACTCAGACAGTGAATTCAGATTCAGC	195541
Db	194402	TTATGCTTTTTCAGTGGCTAAATCTGCTGTTAAAGGTGACATTCACAAATTAACCTGAC	194461	QY	2220	TTCCGAGTTCAGAAATGACTCGGATTCAGATAGCGATTCGACTCAGACAGTGAATTCAGATTCAGC	2279
QY	1140	TAAAGAAATTAACCTTTAAATGGTGTAACTTCAACTGCTAAAGTGCCACCAATTAATGCTGG	1199	Db	195542	TTCCGAGTTCAGAAATGACTCGGATTCAGATAGCGATTCGACTCAGACAGTGAATTCAGATTCAGC	195601
Db	194462	TAAAGAAATTAACCTTTAAATGGTGTAACTTCAACTGCTAAAGTGCCACCAATTAATGCTGG	194521	QY	2280	TAGCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2339
QY	1200	AGATCAAGTATTTGGCAATTTGGTGTAAATCGATAGTGGTAATGTTATTTATACATTTAC	1259	Db	195602	GAGTGAATTCGAGTTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	195661
Db	194522	AGATCAAGTATTTGGCAATTTGGTGTAAATCGATAGTGGTAATGTTATTTATACATTTAC	194581	QY	2340	TTCTGACTTCGCGAGTGAATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2399
QY	1260	AGACTATCTAAATACCTAAAGTGTAAAGTGAATTTGACCACTTGACCACTGGCGCTTATATGA	1319	Db	195662	TTCCGAGTTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	195721
Db	194582	AGACTATCTAAATACCTAAAGTGTAAAGTGAATTTGACCACTTGACCACTGGCGCTTATATGA	194641	QY	2400	CTCAGATAGCGATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2459
QY	1320	CCCTGAAATTTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAC	1379	Db	195722	CTCAGATAGCGATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	195781
Db	194642	CCCTGAAATTTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAC	194701	QY	2460	CAGCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2519
QY	1380	AGCAAAACAAACAGTATTAGTAGATTATGAAATATATCGTAAGTTTATATACTTATCTAT	1439	Db	195782	TAGCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	195841
Db	194702	TGCTAGTAAAGACAGTATTATATCGACTATGAGAAATATGACAAATTTCCATATTTATCAAT	194761	QY	2520	CTCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2579
QY	1440	TAAAGGTACAAATTCGACCAATTCGATAAACAATTAATACGTATCGTACAGCAATTTATGT	1499	Db	195842	CTCAGACTCAGACAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	195901
Db	194762	TAAAGGTACAAATTCGACCAATTCGATAAACAATTAATACGTATCGTACAGCAATTTATGT	194821	QY	2580	CTCAGACTCAGACAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2639
QY	1500	CAATCCAGTGGAGATAACGTTATTGGCGCGTTTAAACAGGTAATTTAAACCAAAATAC	1559	Db	195902	TTCCGATAGTGAATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	195961
Db	194822	CAATCCAGTGGAGATAACGTTATTGGTTACCTTGAACGTAATTTAAACCAAAATAC	194881	QY	2640	CAGTGAATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2699
QY	1560	GGATAGTAAATGATATAGTACAGCAAAATACAGTATTAAGTATATAAGTAGATAA	1619	Db	195962	TAGTGAATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	196021
Db	194882	AAAGAGTAAATGATATAGTACAGCAAAATACAGTATTAAGTATATAAGTAGATAA	194941	QY	2700	TTCCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2759
QY	1620	TGACGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAACTTTGAGATGTCACTAA	1679	Db	196022	TTCCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	196081
Db	194942	TGCTAATGATTTATCTGAAAGTTACTTTGTGAATCCCTAGCGATTTTGAAGATGTAACCTAA	195001	QY	2760	TTCCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2813
QY	1680	TAGTGAATTTATCAATTTCCAAATCCAAATCAATATAAAGTAGAGTTTAAATAGCGCTGA	1739	Db	196082	TTCCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	196141
Db	195002	TCAAGTTAGAAATTTCAATTTCCAAATCCAAATCAATATAAAGTAGAGTTTAAATAGCGCTGA	195061	QY	2814	GAGTGAATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2873
QY	1740	TGATCAATTTACACACCGTATATAGTATGTTTAAATGGTCAATTTGATCCGAATAGCAA	1799	Db	196142	AAAGTGAATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	196201
Db	195062	TGACCAATTTACACACCGTATATAGTATGTTTAAATGGTCAATTTGATCCGATGTAACCTAA	195121	QY	2874	TTCAATAGCGATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2933
QY	1800	AGGTGATTTAGCTTTCACTTTCACTTTATGGGTAAATCTCGATATAATTTGGCGCTC	1859	Db	196202	TTCAATAGCGATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	196261
Db	195122	AGGTGATTTAGCTTTCACTTTCACTTTATGGGTAAATCTCGATATAATTTGGCGCTC	195181	QY	2934	AAATGGTGAATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	2993
QY	1860	TATGTCAATGGGACCAAGTAGCATTTAATACCGATCAGGTTCTGGTACCGGTATCGA	1919	Db	196262	AAATGGTGAATTCGCGAGTGAATTCAGATTCGCGAGTGAATTCGACTCAGACAGTGAATTCGCA	196321
Db	195182	TATGTCAATGGGACCAAGTAGCATTTAATACCGATCAGGTTCTGGTACCGGTATCGA	195241	QY	2994	TACAGGTTCTGAGATGAAGCAATACGTCACCTAAATTTGGGGATTTATAGCATCAATAGG	3053
QY	1920	TAAACAGGTTGTTCTGGAACCACTGATGAGCGTGGTGAATTTGAACCAATTCAGAGGA	1979	Db	196322	TACAGGTTCTGAGATGAAGCAATACGTCACCTAAATTTGGGGATTTATAGCATCAATAGG	196381
Db	195242	TAAACAGGTTGTTCTGGAACCACTGATGAGCGTGGTGAATTTGAACCAATTCAGAGGA	195301	QY	3054	TTCAATTTACTTTTTCAGAGGAAAAAAGAAATAAGATAAGAAATAAGATAATATGAT	3113
QY	1980	TTCCAGATTTCTGACCCAGGTTCCAGATTTCTGCGAGCGATTTCTAATTCAGATAGCGGTTTCA	2039				


```
gene
CDS
DDPKGCGSGSLMNLILQSNFNRHRAIVDKGVLEKACSTLLTTFFKNLRANKKSTN"
4329..5198
/gene="SA0517"
4329..5198
/gene="SA0517"
/note="ORFID:SA0517"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB41748.1"
/db_xref="GI:13700451"
/translacion="MKLIATDMGTLNAAHEISQPNIDAIKYAQEQIGTVIATGR
AFYERAAQVATDLTVPICLNGAEVROETFMVSTSHLNSLVHKIINVLKQAGIYY
QYVIRAITSDPDLDIYIDIAERAGOHANVERIKSGIQRIDNGLTKVVDNYDAI
ENIPELIMKLAFDNELEKIDKASKILAEFNPALIKSSSRGNIEITHSDAKGTALE
TIAERLIGLEMDKDVMAIGNDLMSLEKGVGVPMANGAEVKKIAKVVDINENSGV
GKAIMKLIREQOV"
5220..5786
/gene="SA0518"
5220..5786
/gene="SA0518"
/note="ORFID:SA0518"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB41749.1"
/db_xref="GI:13700452"
/translacion="MKGLIILIGSQVNSHTSALARYLPEHFKTHDIEAEIPDLAEKPL
NQLDFSGTPTSIDIKOMKDLKEMAAADPLILGTENYHGSYSGILKALDHLNMDY
FKMPVGLIGSGGIVSESEPLSHLRVIVRSLLIGTAVPQIATHDSDFAKNEDGGYYLN
DSEFQLRLARLEVDQIVSEFVNSPYEHLK"
6215..9076
/gene="sdrC"
6215..9076
/gene="sdrC"
/note="ORFID:SA0519"
/codon_start=1
/transl_table=11
/product="Ser-Asp rich fibrinogen-binding, bone
sialoprotein-binding protein"
/protein_id="BAB41750.1"
/db_xref="GI:13700453"
/translacion="MNNKKTATNRKGMIPNRLNFKSRKYSVGTASILVGTTLIFGUS
GHEKAAHNGELNQSNETTAPSENKTTIEKVDRLQKNDKNTQATADQPKVMSDSA
TVKETSNNQPNATASQSTQTSNVTTNDKSSITYSNEDKSNLTQAKNVSTPTKT
TKIKRALNRMAVNTAAPQGTNVNDKVFHTNIDIAIDKHVNKTTGNTFEFATSSD
VLKLANYTIDDSVKEGDTTFKYGOVFRPGSVRLPSQTNLYNAQGNIIAKGLYDSK
TNTTYFTNVVDQYTNVSGSFEOVAFKSENATTDKATXMEVTLGNDIYSKDVIVD
YGNQKQQLISSINYINNELSRNMTVYNQPKTYTKETFTVNTLITGYKENPDANKPK
IYEVTDNQFVDSFTPTSKLDVITGQFVYISNDKNTATVDLLNGSSDRQYIIQQ
VAYPDSNTDNGKIDITYLETQNGKSSNSYSNVNGSSTANGQPKYKYNLGDYVWEDTN
KDGQDANEKGIKGVYVILKDSNGKELDRITTDENGKYQFTGLSNGYSVEFSPAGY
TPTTANAGTDVDSGLTGTGTGVDADNMTLDSGFYKTPKYSILGDYVWVDSNKGQK
DSTEKGIKGVKVLQNEKGVIGTETDENGKVFEDNLDGSKYVIEKPAGLTQGTG
NTTEDDKADQGVVDVITDHDFTLNGYVEETSDDSDSDSDSDSDSDSDSDSDSDSD
DSUSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD
DSUSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD
ALPETGSENNNSNGTLFGGLFAALGSLILFGRKKQNK"
9443..13600
/gene="sdrD"
9443..13600
/gene="sdrD"
/note="ORFID:SA0520"
/codon_start=1
/transl_table=11
/product="Ser-Asp rich fibrinogen-binding, bone
sialoprotein-binding protein"
/protein_id="BAB41751.1"
/db_xref="GI:13700454"
/translacion="MLNRENKTAITRGMVSNRLNFKFSIKYTVGTASILVGTTLIFG
LGNEAKAEASTNKELENTTASDQSSDKVDMQOLNQEDNTKDNOKEMVSSQNE
TTSNGNKSIEKESVQSTTGKNEVSTAKSDEQASPKSTNEDLNTKQTSINQEGLPDL
```

```
Query Match 81.6%; Score 2853.4; DB 1; Length 299050;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 3232; Conservative 0; Mismatches 266; Indels 169; Gaps 5;

QY 1 GGTACATAAATTACACATCTGCTTTTGAAGAAATATGATTTCAAGCTAGGATTACATTA 60
DB 249183 GGTACCATTAATTACACATCTGCTTTTGAAGAAATATGATTTTAAAGCTAGGATTACATTA 249242

QY 61 GGTAGAGTTCATTAATAATAAATAAATGTTTGCATCAAAATCGTACCTTGTGCTTGTGA 120
DB 249243 GGTAGAGTTCATTAATAATAAATAAATGTTTGCATCAAAATCGTACCTTGTGCTTGTGA 249302

QY 121 ATTCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTAATTGTGGAATAATAAAA 180
DB 249303 ATTCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTAATTGTGGAATAATAAAA 249362

QY 181 TATCGATACAAATTAATTGCTATTAATGCAATTTTAGTGCTATAAATTCATTACAGAGATT 240
DB 249363 TATCGATACAAATTAATTGCTATTAATGCAATTTTAGTGCTATAAATTCATTACAGAGATT 249422

QY 241 AAATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTTAAAGAGAGGGAATAA 299
DB 249423 AAATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTTAAAGAGAGGGAATAA 249482

QY 300 AATCAATATGAGAAAGAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 359
DB 249483 AATCAATATGAGAAAGAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 249542

QY 360 AGTCCTTGTAGTAGTACGTTAATCGGTTTTTGGACTACTCAGCAGTAAAGAGAGAGATGCAAG 419
DB 249543 AGTCCTTGTAGTAGTACGTTAATCGGTTTTTGGACTACTCAGCAGTAAAGAGAGAGATGCAAG 249602

QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAAATGATTCAAG 479
DB 249603 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAAATGATTCAAG 249662

QY 480 TAGCGTTAGTGTGCTGCACTTAAACAGACAGACACAAACGTAAGTGATATCTAAACATCGTC 539
DB 249663 TAGCGTTAGTGTGCTGCACTTAAACAGACAGACACAAACGTAAGTGATATCTAAACATCGTC 249722

QY 540 AAACACTATAAATAGCGGAAACGAGTGTGGCGCAAAATCCAGCACAACAGAGAAACGACACA 599
DB 249723 AAACACTATAAATAGCGGAAACGAGTGTGGCGCAAAATCCAGCACAACAGAGAAACGACACA 249782

QY 600 ATCATCATCAAAATGCAACTACGGAAGAAACGCCGGTAACCTGGTGAAGCTACTACTAC 659
DB 249783 ATCATCATCAAAATGCAACTACGGAAGAAACGCCGGTAACCTGGTGAAGCTACTACTAC 249842

QY 660 GACACGAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGA 719
DB 249843 GACACGAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGA 249902

QY 720 ATTAGTGAATCAAAACAAAGTAATGAAACGACTTTTAAATGATCTACTAATACAGTATCATCTGT 779
DB 249903 ATTAGTGAATCAAAACAAAGTAATGAAACGACTTTTAAATGATCTACTAATACAGTATCATCTGT 249962

QY 780 AATTTCACCTCAAAATCTCAAAATCGGGAATAATGTTTCAACCAACCGAAGATCTTCAAC 839
```

Db	249963	AAATTCCACTCCAAATTTCTCAAAATGCGGAAAATGTTTCAACCAACGACGATACACTTCAAC	250022
Qy	840	TGAAGCAAACACCTTCAAAACAATGAATCAGCTCCACAGAGTACAGATGCAAGTAAATAAGA	899
Db	250023	TGAAGCAACACCTTCAAAACAATGATCAGCTCCACAGAAATACAGATGCAAGTAAATAAGA	250082
Qy	900	TGTAGTTAAATCAAGCGGTTAAATCAAGTGGCGCTAGAAATGAGAGCAATTTAGTTTAGCGGC	959
Db	250083	TGTAGTTAGTCAAGCGGTTAAATCAAGTACGCTAGAAATGAGAGCAATTTAGTTTAGCGGC	250142
Qy	960	AGTAGCTGCAGATGCAACCGCAGCTGCGACAGATATTTACGAATCAGTTGACGAATGTGAC	1019
Db	250143	AGTAGCTGCAGATGCAACCGCAGCTGCGACAGATATTTACGAATCAGTTGACGAATGTGAA	250202
Qy	1020	AGTTGGTATTGACTCTGCTGACGATCTGTGTATCGCACCAAGCAGGTTATGTCACAACTGAA	1079
Db	250203	AGTTTACTATTGACTCTGCTGACGATCTGTGTATCGCACCAAGCAGGTTATGTCACAACTGAA	250262
Qy	1080	TTATGGTTTTTTCAGTGCGCTAATTTCTGCTGTTTAAAGGTGACACATTTCAAAATTAAGTACC	1139
Db	250263	TTATGGTTTTTTCAGTGCGCTAATTTCTGCTGTTTAAAGGTGACACATTTCAAAATTAAGTACC	250322
Qy	1140	TAAAGAATTAACCTTAAATGGTGTAACCTTCAACTGCTAAAGTGCCACCAATTTATGGCTGG	1199
Db	250323	TAAAGAATTAACCTTAAATGGTGTAACCTTCAACTGCTAAAGTGCCACCAATTTATGGCTGG	250382
Qy	1200	AGATCAAGTATTGGCAATGCTGAATCGATAGTGATGGTAATGTTTATATACATTTAC	1259
Db	250383	AGATCAAGTATTGGCAATGCTGAATCGATAGTGATGGTAATGTTTATATACATTTAC	250442
Qy	1260	AGACTATGTAATATCTAAAGATGATGTAAGAACCTTTGACCAATGCCGCTTATATTGA	1319
Db	250443	AGACTATGCTGATATAAAGAAAAATGTAAACAGCTAAATATTACTATGCCAGCTTATATTGA	250502
Qy	1320	CCCTGAAAATGTTAAAAAGACAGGTAATGTGCATTTGGCTACTCGCATPAGTGTACAAAC	1379
Db	250503	CCCTGAAAATGTTCAAAAGACAGGTAATGTGCATTTGACATTCACACTGGCATPAGAAACCAATAC	250562
Qy	1380	AGCAAAACAAAACAGTATTAGTATGATTATGAAAATATCGTAAAGTTTTTATAACTTATCTAT	1439
Db	250563	TGCTAGTAAACAGTATTAAATCGACTATGAGAAATATGCAAAATTTCCATAATTTATCAAT	250622
Qy	1440	TAAAGGTACAAATTGACCAATTCGATPAAACAAATAATACGTATCGTTCAGACAAATTTATGT	1499
Db	250623	TAAAGGTACGATTTGATCAAAATCGATAAAACAAATAATACGTATCGCCTCCAAACAAATTTATGT	250682
Qy	1500	CAATCCAGTGGAGATACGTTTATTTGCGCGGTTTTTACACGGTAATTTTAAACCAAAATAC	1559
Db	250683	CAATCCACGCGGAGATAAGTTGTGTACCTGCCCTTACACGGTAATTTTAAATCCCTAATAC	250742
Qy	1560	GGATAGTAATGCATTTAAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAA	1619
Db	250743	AAAGAGTAATGCGTTTAAATAGATGCAAAACAACTGATTTAAAGTTTATAGAGTCGATAA	250802
Qy	1620	TGCAGCTGATTTATCTGAAAGTTACTTTGTGTAATCCAGAAAACCTTTGAGGATGTCACCTAA	1679
Db	250803	TGCTAATGATTTATCTGAAAGTTATTTATGTGAATCCTAGCGATTTTGAAGATGTAACATAA	250862
Qy	1680	TAGTGTGAATATTAATCCCAAAATCCAAATCAATATAAAGTAGAGTTTAAATACCCTCTGA	1739
Db	250863	TCAAGTTAGAAATTTCCATTTCCAAATGCTAATCAATACAAAGTAGAATTTTCTTACGACGA	250922
Qy	1740	TGATCAAAATTAACAACCGTATATAGTGTGTTAATGCTCATTTATGTATCCGAATAGCAA	1799
Db	250923	TGACCAAAATTAACAACCGTATATAGTGTGTTAATGCGCATTTGATCTCTGCTAGTAC	250982
Qy	1800	AGGTGATTTAGCTTTTACGTTTCACTTTATATGGGTATACTCGAATATATATTTGGCGCTC	1859
Db	250983	AGGTGATTTAGCACTACGTTTCGACATTTATGGTTATGATTTCTAAATTTATATGGAGATC	251042
Qy	1860	TATGTCATGGGACAAACGAAGTAGCATTTTAAATAACGGATCAGGTTCTGGGTACGGTATCGA	1919
Db	251043	TATGTCATGGGACAAACGAAGTAGCATTTTAAATAACGGATCAGGTTCTGGGTACGGTATCGA	251102

QY	1920	TAAACAGTTGTTCTTCACACACCTCATGAGCTGGTGGAAATTGAACCAATTCCAGAGGA	1979
Db	251103	TAAACAGTTGTTCTTCGAAACAACCTGATGAGCTGGTGGAAATTGAACCAATTCCAGAGGA	251162
QY	1980	TTCAGATTCTGTACCCAGGTTTCAGATTCTGCGACGGATTCTTAATTTCAGATAGCGGTTCCAGA	2039
Db	251163	TTCAGATTCTGTACCCAGGTTTCAGATTCTGCGACGGATTCTTAATTTCAGATAGCGGTTCCAGA	251222
QY	2040	TTCCGGTAGTGATTTCTACATCAGATAGTGTTTCAGATTTCAGCGGAGTATTTCAGATTCCAGC	2099
Db	251223	TTCTGCGCAGTGATTTCTACATTCAGATAGTGTTTCAGATTTCAGCGGAGTATTTCAGATTCCAGC	251282
QY	2100	AACTGATTTCAGATCTCAGCGAGTGAATTCAGATTTCAGCA-----	2136
Db	251283	AACTGATTTCAGATCTCAGCGAGTGAATTCAGATTTCAGCAAGTGAATTCAGATTTCAGCAAGTGA	251342
QY	2137	-----	2136
Db	251343	TTCCAGATTTCAGCAAGTGATTTTCAGACTTCAGCAAGTGATTTTCAGATTTCAGCAAGTGATTTTCAGA	251402
QY	2137	-----ACCGATTCCGACTTCAGCGAGCGAATTCGGAGCTTCAGA	2171
Db	251403	TTCCAGCAAGCGAATTCAGATTTCAGCGAGCGAATTCAGATTTCAGCGAGCGAATTCAGATTTCAGC	251462
QY	2172	CAATGACTTCGGATTTCAGATAGCGAATTCCTGACTTCAGACAGTGACTTCAGATTTCGGACAGTGA	2231
Db	251463	GAGTGATTCCGACTTCAGCGAGCGAATTCAGACTTCAGATAGTGACTTCAGATTTCGGATAGCGGA	251522
QY	2232	CTCAGATTTCAGATAGCGAATTCCTGACTTCAGACAGTGACTTCAGATTTCAGATAGCGAATTCAGA	2291
Db	251523	TTCCGACTTCAGATAGCGAATTCAGATTTCGGATAGCGAATTCCTGACTTCAGACAGCGAATTCAGA	251582
QY	2292	TTCCAGATAGCGAATTCAGATTTCGGACTTCGGACTTCAGACAGCGAATTCCTGACTTCGACTCCGA	2351
Db	251583	CTCAGACAGTGACTTCAGATTTCGGATAGCGAATTCCTGACTTCAGACAGCGAATTCAGATTTCGGA	251642
QY	2352	CAGTGATTTCGACTTCAGAC-----AGCGATTTCAGATTTCGGACAGTGA	2393
Db	251643	TAGCGATTTCGACTTCAGACAGTGACTTCAGATTTCGGATAGCGAATTCGGATTCGGACAGTGA	251702
QY	2394	TTCCGACTTCAGATAGCGAATTCGGACTTCAGATAGCGAATTCAGATTTCAGACAGCGAATTCAGA	2453
Db	251703	TTCCGACTTCAGATAGCGAATTCGGACTTCAGATAGCGAATTCAGATTTCAGACAGCGAATTCAGA	251762
QY	2454	TTCCAGACAGCGAATTCAGATTTCAGATAGCGAATTCAGATTTCGGACAGTGACTTCAGATTTCGGA	2513
Db	251763	TTCCAGACAGCGAATTCGACTTCAGACAGTGACTTCAGATTTCGGATAGCGAATTCAGATTTCAGA	251822
QY	2514	CAGTGACTTCGGATTTCAGATAGCGAATTCAGATTTCGGACAGTGACTTCAGATTTCGGACAGTGA	2573
Db	251823	CAGTGATTTCAGACTTCAGATAGCGAATTCAGATTTCGGACAGTGACTTCAGATTTCGGACAGTGA	251882
QY	2574	CTCAGACTTCAGACAGTGATTTCGGATTTCAGCGAGTGAATTCGGATTTCAGATAGTGATTTCGGA	2633
Db	251893	TTCCAGATTTCGGATAGCGAATTCAGATTTCGGACAGTGACTTCAGATTTCGGATAGTGATTTCGGA	251942
QY	2634	TTCCGACAGTGACTTCGGATTTCAGATAGCGAATTCAGACTTCGGATAGCGAATTCGGATTTCAGA	2693
Db	251943	TTCCGCGAGTGATTTCAGATTTCAGATAGCGAATTCAGATTTCGGATAGTGACTTCAGACTTCAGA	252002
QY	2694	TAGCGAATTCGGATTTCAGATAGCGAATTCAGATTTCGGACAGTGATTTCAGATTTCAGACAGCGA	2753
Db	252003	CAGTGATTTCAGATTTCAGATAGTGACTTCAGATTTCAGACAGCGAATTCAGATTTCAGATAGTGGA	252062
QY	2754	TTCCAGATTTCAGACAGC-----GACTTCAGACAGTGACTTCAGATTTCAGATAGTGACTTCGGA	2807
Db	252063	CTCCGATTTCAGACAGCGAATTCAGATTTCAGATTTCAGATTTCGGATTTCAGATAGTGATTTCGGA	252122
QY	2808	TTCCAGCGAGTGATTTCAGACTTCAGGTAGTGACTTCGGATTTCATCAAGTGATTTC	2858
Db	252123	TTCCAGCGAGTGATTTCAGACTTCAGGTAGTGACTTCGGATTTCATCAAGTGATTTCAGATTTCGGA	252182

301 ATGAATATGAAGAAAAAGAAAAACACGCAATTCGGAAAAAATCGAATTCGGCTGGCTTCA 360
Db
1 ATGAATATGAAGAAAAAGAAAAACACGCAATTCGGAAAAAATCGAATTCGGCTGGCTTCA 60
361 GTGCTGTAGAGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAGAAGCAGATCGAAGT 420
Db
61 GTGCTGTAGAGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAGAAGCAGATCGAAGT 120
421 GAAAAATAGTTTACCAATCTCATAGCGCAAGTAAAGCAAGCAAGTAATGATTCAGT 480
Db
121 GAAATAGTTTACCAATCTCATAGCGCAAGTAAAGCAAGCAAGTAATGATTCAGT 180
481 AGCGTTAGTGTGACCTTAAACAGACGACACAAACGTCAGTGAATCAAAAAACATCGTCA 540
Db
181 AGCGTTAGTGTGACCTTAAACAGACGACACAAACGTCAGTGAATCAAAAAACATCGTCA 240
541 AACCTAATATGCGGAACGAGTGTGGGCHAAATCCAGCAACAGAAACGACACAA 600
Db
241 AACCTAATATGCGGAACGAGTGTGGGCHAAATCCAGCAACAGAAACGACACAA 300
601 TCATCATCAAAATCAACTACGGAAGAAACGCGGTAACCTGTTGAAGCTACTACTAG 660
Db
301 TCATCATCAAAATCAACTACGGAAGAAACGCGGTAACCTGTTGAAGCTACTACTAG 360
661 ACAAGCAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGCGAGGAA 720
Db
361 ACAAGCAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGCGAGGAA 420
721 TTAGTGAAATCAAGTAATGAACGACTTTAATGATCTAATGATCTAATGATCTAATGATCTA 780
Db
421 TTAGTGAAATCAAGTAATGAACGACTTTAATGATCTAATGATCTAATGATCTAATGATCTA 480
781 AATTCACTCAAAATCTTACAAATCGGAAATGTTTCAACAAACGCAAGACTTCAACT 840
Db
481 AATTCACTCAAAATCTTACAAATCGGAAATGTTTCAACAAACGCAAGACTTCAACT 540
901 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT 600
Db
541 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT 600
901 GTAGTTAATCAAGCGTTAATACAGTCCGCTAGATGAGAGCAATTTAGTTTACGGGCA 960
Db
601 GTAGTTAATCAAGCGTTAATACAGTCCGCTAGATGAGAGCAATTTAGTTTACGGGCA 660
961 GTAGCTCAGATGACCGGAGCTGGCACAGATATACGAATCAGTTGACGAATGTGACA 1020
Db
661 GTAGCTCAGATGACCGGAGCTGGCACAGATATACGAATCAGTTGACGAATGTGACA 720
1021 GTTGGTATGACTCTGGTACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAACTGAAT 1080
Db
721 GTTGGTATGACTCTGGTACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAACTGAAT 780
1081 TATGTTTTTTCAGTGCCTAATCTGCTGTTAAAGGTGACACATTCAAATTAACGTACCT 1140
Db
781 TATGTTTTTTCAGTGCCTAATCTGCTGTTAAAGGTGACACATTCAAATTAACGTACCT 840
1141 AAAGAAATTAACCTTAAATGTTAACTTCAACTGCTAAAGTGCCACCAATTTATGGCTGGA 1200
Db
841 AAAGAAATTAACCTTAAATGTTAACTTCAACTGCTAAAGTGCCACCAATTTATGGCTGGA 900
1201 GATCAAGTATGCGAAATGTTAATCGATAGTGATGGTAATGTTATTTATACATTTACA 1260
Db
901 GATCAAGTATGCGAAATGTTAATCGATAGTGATGGTAATGTTATTTATACATTTACA 960
1261 GACTATCTAAATACTAAAGATGATTAAGCACTTTGACCAATTCGCGCTTATTTGAC 1320
Db
961 GACTATCTAAATACTAAAGATGATTAAGCACTTTGACCAATTCGCGCTTATTTGAC 1020
1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGCTACTGTCATAGGTAGTACAAACA 1380
Db
1021 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGCTACTGTCATAGGTAGTACAAACA 1080

1381 GCAACCAAAACAGTATTAGTAGATTGAAAAATATGTAAGTTTATTAACCTATCTATT 1440
Db
1081 GCAACCAAAACAGTATTAGTAGATTGAAAAATATGTAAGTTTATTAACCTATCTATT 1140
1441 AAAGTCACAATTTGACCAAAATCGATAAACAACAAATAATACGTATCGTCAGACAAATTTATGTC 1500
Db
1141 AAAGTCACAATTTGACCAAAATCGATAAACAACAAATAATACGTATCGTCAGACAAATTTATGTC 1200
1501 AATCAAGTGAGATGAAGTTTATTCGCGGGGTTTTAAACAGGTAATTTAAAAACCAATACG 1560
Db
1201 AATCAAGTGAGATGAAGTTTATTCGCGGGGTTTTAAACAGGTAATTTAAAAACCAATACG 1260
1561 GATAGTAATGCAATTAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAAT 1620
Db
1261 GATAGTAATGCAATTAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAAT 1320
1621 GCAGCTGATTTATCTGAAAAGTTTACTTTTGTGAATCCAGAAAACTTTGAGGATCTCACTAAT 1680
Db
1321 GCAGCTGATTTATCTGAAAAGTTTACTTTTGTGAATCCAGAAAACTTTGAGGATCTCACTAAT 1380
1681 AGTGTGAATATTACATTCGCAAAATCCAAATCAATATAAAGTAGAGTTTAAATACGCTGAT 1740
Db
1381 AGTGTGAATATTACATTCGCAAAATCCAAATCAATATAAAGTAGAGTTTAAATACGCTGAT 1440
1741 GATCAAAATTACAAACACCGTATATAGTAGTTTGTAAATGCTCATATTTGATCCGAAATAGCAAA 1800
Db
1441 GATCAAAATTACAAACACCGTATATAGTAGTTTGTAAATGCTCATATTTGATCCGAAATAGCAAA 1500
1801 GGTGATTTAGCTTTTACGTTCAACTTTATATGGGTATAACTCGAAATATAATTTGGCGCTCT 1860
Db
1501 GGTGATTTAGCTTTTACGTTCAACTTTATATGGGTATAACTCGAAATATAATTTGGCGCTCT 1560
1861 ATGTCATGGGACACGAGTAGCAATTTAATAACCGATCAGGTTCTGTTGAGCGGTATCGAT 1920
Db
1561 ATGTCATGGGACACGAGTAGCAATTTAATAACCGATCAGGTTCTGTTGAGCGGTATCGAT 1620
1921 AAACAGTTGTTCTCTGAACCACTGATGAGCTCGTGAAATTTGAACCAATTTCCAGAGAT 1980
Db
1621 AAACAGTTGTTCTCTGAACCACTGATGAGCTCGTGAAATTTGAACCAATTTCCAGAGAT 1680
1981 TCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTCTAATTCAGATAGCGTTTCAGAT 2040
Db
1681 TCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTCTAATTCAGATAGCGTTTCAGAT 1740
2041 TCGGGTAGTGTATCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGGTTTCAGATTCAGCA 2100
Db
1741 TCGGGTAGTGTATCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGGTTTCAGATTCAGCA 1800
2101 AGTGATTCAGACTCAGCGAGTGGTTTCAGATTCAGCGAGTGGTTTCAGACTCAGCGAGGAT 2160
Db
1801 AGTGATTCAGACTCAGCGAGTGGTTTCAGATTCAGCGAGTGGTTTCAGACTCAGCGAGGAT 1860
2161 TCCGACTCAGACAAATGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTGAATCCAGAT 2220
Db
1861 TCCGACTCAGACAAATGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTGAATCCAGAT 1920
2221 TCCGACTCAGACAAATGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTGAATCCAGAT 2280
Db
1921 TCCGACTCAGACAAATGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTGAATCCAGAT 1980
2281 AGCGATTCAGATTCAGATAGCGATTTCAGATTCGCGAGTGGTTTCGAGCTCAGACAGCGAT 2340
Db
1981 AGCGATTCAGATTCAGATAGCGATTTCAGATTCGCGAGTGGTTTCGAGCTCAGACAGCGAT 2040
2341 TCTGACTCCGACAGTGGTTCCGACTCAGACAGCGATTTCAGATTCGCGAGTGGTTCCGAC 2400
Db
2041 TCTGACTCCGACAGTGGTTCCGACTCAGACAGCGATTTCAGATTCGCGAGTGGTTCCGAC 2100
2401 TCAGATAGCGATTCGCGACTCAGATAGCGATTTCAGATTCGCGAGTGGTTCCGACTCAGAC 2460
Db
2101 TCAGATAGCGATTCGCGACTCAGATAGCGATTTCAGATTCGCGAGTGGTTCCGACTCAGAC 2160
2461 AGCGATTCAGATTCAGATAGCGATTTCAGATTCGCGAGTGGTTCCGACTCAGATTCGCGAGT 2520

2161	Db	AGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGACAGTGA	2220
2521	Qy	TCGGATTTCAGATAGCGATTTCAGATTTCGACAGTGA	2580
2221	Db	TCGGATTTCAGATAGCGATTTCAGATTTCGACAGTGA	2280
2581	Qy	TCAGACAGTGA	2640
2281	Db	TCAGACAGTGA	2340
2641	Qy	AGTGACTTCGGATTTCAGATAGCGACTTCAGATTTCGAGATAGCGAT	2700
2341	Db	AGTGACTTCGGATTTCAGATAGCGACTTCAGATTTCGAGATAGCGAT	2400
2701	Qy	TCGGACTTCAGATAGCGATTTCAGATTTCGACAGCGATTTCAGATTTCAGAT	2760
2401	Db	TCGGACTTCAGATAGCGATTTCAGATTTCGACAGCGATTTCAGATTTCAGAT	2460
2761	Qy	TCAGACAGCGACTTCAGACAGTGA	2820
2443	Db	TCAGACAGCGACTTCAGACAGTGA	2502
2821	Qy	TCAGACTTCAGATAGTGA	2880
2503	Db	TCAGACTTCAGATAGTGA	2562
2881	Qy	AGCGATTTCGGAGTTCAGATTTCGAGATAGCGATTTCGAGATAGCGAT	2940
2563	Db	AGCGATTTCGGAGTTCAGATTTCGAGATAGCGATTTCGAGATAGCGAT	2622
2941	Qy	ACTAATGCTTCTAATAAATAGAGGCTTAAGATAGTAAAGAACCAATTACAGATACAGGT	3000
2623	Db	ACTAATGCTTCTAATAAATAGAGGCTTAAGATAGTAAAGAACCAATTACAGATACAGGT	2682
3001	Qy	TCTCAAGATGAAGCAAAATACGTCACTAATTTCGGGATTAATTAGCATCAATAGTGTTTCATTA	3060
2683	Db	TCTCAAGATGAAGCAAAATACGTCACTAATTTCGGGATTAATTAGCATCAATAGTGTTTCATTA	2742
3061	Qy	CTACTTTTTCAGAGAAAAAAGAAAAATTAAGATTAAGAAA	3099
2743	Db	CTACTTTTTCAGAGAAAAAAGAAAAATTAAGATTAAGAAA	2781

RESULT 12

RESULI 12
RX571856 08

BY 3/18/50
WPCOMMENT

COMMENT
Sequence split into 29 fragments LOCUS BX571856 Accession BX571856

Fragment Name	Begin	End
1	110000	
BX571856_00	100001	210000
BX571856_01	200001	310000
BX571856_02	300001	410000
BX571856_03	400001	510000
BX571856_04	500001	610000
BX571856_05	600001	710000
BX571856_06	700001	810000
BX571856_07	800001	910000
BX571856_08	900001	1010000
BX571856_09	1000001	1110000
BX571856_10	1100001	1210000
BX571856_11	1200001	1310000
BX571856_12	1300001	1410000
BX571856_13	1400001	1510000
BX571856_14	1500001	1610000
BX571856_15	1600001	1710000
BX571856_16	1700001	1810000
BX571856_17	1800001	1910000
BX571856_18	1900001	2010000
BX571856_19	2000001	2110000
BX571856_20	2100001	2210000
BX571856_21	2200001	2310000
BX571856_22	2300001	2410000
BX571856_23	2400001	2510000

Db	88837	ATGTTAGTTAAATCAAGCGGTTAATCAAGTCGCGCTAGAAATAGAGAGCAATTTAGTTTAGCGG	88896
Qy	959	CAGTAGCTCAGATGCACCGGAGCTGGCAGAGATTAACGAATACGAATCAGTTGACGAATGTA	1018
Db	88897	CTGTAGCTCAGATGCACCGGCTGTGGCAAGATATTACGAATCAGTTGACGAATGTA	88956
Qy	1019	CAGTTGGTATTGACTCTGTGACGAGTGTATCCGACCAAGCAGGTTATGTCAAACTGA	1078
Db	88957	CAGTTGGTATTGACTCTGTGAGATACAGTTTATCCGACCAAGCAGGCTATGTCAAACTGA	89016
Qy	1079	ATTATGGTATTTCAGTGGCTTAATCTCTGCTTTAAAGGTGACACATTCAAATTAACGTAC	1138
Db	89017	ATTATGGGTTCTCAGTACCAATAGGCGCTGTCAAGGTGACACATTCAAATTAACGTGC	89076
Qy	1139	CTAAAGAAATTAACCTTAATGGTGTAACCTTCAACTGCTAAAGTGCCACCAATTTAGGCTG	1198
Db	89077	CCAAAGAAATTAACCTTAATGGTGTAACCTTCAACTGCTAAAGTGCCACCAATTTAGGCGG	89136
Qy	1199	GAGATCAAGTATTGGCAATGGTGAATCGATAGTGAATGTTATTTATATCAATTTA	1258
Db	89137	GAGATCAAGTATTGGCAATGGTGAATCGATAGTGAATGTTATTTATATCAATTTA	89196
Qy	1259	CAGACTATGTAATTAACGATGATGTAAGCAACTTTGACCAATGCCCGTTTATATG	1318
Db	89197	CAGACTATGTAATTAACGATGATGTAAGCAACTTTGACCAATGCCCGTTTATATG	89256
Qy	1319	ACCTGAAATCTTAAAGAGCAGGTAATGTGACATTCGCTACTGGCATAGTAGTACAA	1378
Db	89257	ACCTGAAATCTTAAAGAGCAGGTAATGTGACATTCGCTACTGGCATAGTAGTACAA	89316
Qy	1379	CAGCAACAAACAGTATTAGTATGTAAGTAAATATGGTAAGTTTATTAACCTTATCTA	1438
Db	89317	CAGCAACAAACAGTATTAGTATGTAAGTAAATATGGTAAGTTTATTAACCTTATCTA	89376
Qy	1439	TTAAAGGTACAAATTCGCAATTCGATTAACCAATTAACGTAACGTTTATG	1498
Db	89377	TTAAAGGTACAAATTCGCAATTCGATTAACCAATTAACGTAACGTTTATG	89436
Qy	1499	TCAATCCAAAGTGGAGATAACGTTATTCGCGCGGTTTAAACAGGTAATTTAAACCAATA	1558
Db	89437	TCAATCCAAAGTGGAGATAACGTTATTCGCGCGGTTTAAACAGGTAATTTAAACCAATA	89496
Qy	1559	CGGATAGTAATGCAATTAATAGATCAGCAAAATACAAGTATTAAGTATATAAGTAGATA	1618
Db	89497	CGGATAGTAATGCAATTAATAGATCAGCAAAATACAAGTATTAAGTATATAAGTAGATA	89556
Qy	1619	ATGCAGCTGATTTATCTGAAAGTTACTTGTGATCCAGAAACTTTTCAGGATGTCATA	1678
Db	89557	ATGCATCAGACTTGTCTGAAAGTTATTATGTAATCCAGATAAATTTGAAGATGTCATG	89616
Qy	1679	ATAGTGTGAATTTATACATTCCTCAATCCAAATCAATCAATTAAGTAGAGTTTAAACGCTG	1738
Db	89617	ATAGTGTGAATTTATACATTCCTCAATCCAAATCAATCAATTAAGTAGAGTTTAAACGCTG	89676
Qy	1739	ATGATCAAAATTAACACCGGTATATAGTGTGTAATAGTATATTTGATCCGAATAGCA	1798
Db	89677	ATGATCAAAATTAACACCGGTATATAGTGTGTAATAGTATATTTGATCCGAATAGCA	89736
Qy	1799	AAGTGTATTTAGCTTTACGTTCACTTTATATGGTATTAACCTGCAATTAATTTGGCGCT	1858
Db	89737	AAGTGTATTTAGCTTTACGTTCACTTTATATGGTATTAACCTGCAATTAATTTGGCGAT	89796
Qy	1859	CTATGTCATGGGCAACGAAGTAGCAATTAATTAACCGATCAGGTTCTGTTGACGGTATCG	1918
Db	89797	CAATGTCATGGGCAATGAAGTAGCAATTAATTAACCGATCAGGTTCTGTTGACGGTATCG	89856
Qy	1919	ATAAACGAGTTGTCCTGAACACCTGATGAGCTGTGAAATTAACCAATTCAGAGG	1978
Db	89857	ATAAACGAGTTGTCCTGAACACCTGATGAGCTGTGAAATTAACCAATTCAGAGG	89916
Qy	1979	ATTACAGATTCGACCCCA- -----GGTTCAGATTCGCGAGGATTTCTAAATTCAG	2026
Db	89917	ATTACAGATTCGACCCAGGTTTCAGATAGTGGTTCAGATTCGCGAGGATTTCTAAATTCAG	89976
Qy	2027	ATAGCGGTTTCAGATTCGGGTAGTGTATTCTACATCAGATAGTGGTTCAGATTCAGCGAGTG	2086
Db	89977	ATAGCGGTTTCAGATTCGGGTAGTGTATTCTACATCAGATAGTGGTTCAGATTCAGCGAGTG	90036
Qy	2087	ATTTCAGATTCAGCAAGTGTATTTCAGACTCAGCGAGTGTATTTCAGATTCAGCAAGCGATTCCG	2146
Db	90037	ATTTCAGATTCAGCAAGTGTATTTCAGACTCAGCGAGTGTATTTCAGATTCAGCAAGCGATTCCG	90096
Qy	2147	ACTCAGCGAGCGAT-----	2160
Db	90097	ATTTCAGCGAGCGATTCAGATTCAGCAAGTGTATTTCAGACTCAGCGAGTGTATTTCAGATTCAA	90156
Qy	2161	-----	2160
Db	90157	CGAGTGTATTCAGATTCAGCAAGTGTATTTCAGACTCAGCAAGCGATTCAGATTCAGATAGTG	90216
Qy	2161	--TCGACTCAGACAAATGACTTCGGATTCAGATAGCGATTCGACTCAGACAGTGTACTCAG	2218
Db	90217	ACTCAGACTCAGACAGCGATTCAGATTCAGATAGTGTACTCAGACTCAGATAGCGATTCAG	90276
Qy	2219	ATTTCGACAGTGTACTCAGATTCAGATAGCGATTCGACTCAGACAGTGTACTCAGATTCCAG	2278
Db	90277	ATTTCGACAGTGTACTTCGGATTCAGATAGCGATTCGACTCAGACAGTGTACTCAGATTCCAG	90336
Qy	2279	ATAGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACAGTGTATTC-----	2325
Db	90337	ATAGCGACTCAGACTCAGAAAGCGACTCAGATTCAGACAGCGATTCGCGACTCAGATAGTG	90396
Qy	2326	-----	2325
Db	90397	ATTTCAGATTCAGACAGCGACTCAGACAGTGTATTTCAGATTCAGACAGTGTACTCAGACTCGG	90456
Qy	2326	-----	2325
Db	90457	ATAGTGAATCAGACTCAGACAGCGACTCAGATTCAGACAGCGACTCAGACTCGGATAGTG	90516
Qy	2326	-----GACTCAGACAGCGATTCGACTTCGACAGTGTATTTCGACTCAGACAGCGATTCAG	2380
Db	90517	AATCAGACTCAGACAGTGTACTCAGACTCAGACAGCGACTCAGACTCAGATAGCGACTCAG	90576
Qy	2381	ATTCGCGAGTGTATTCGACTCAGATAGCGATTCGCGACTCAGATAGCGACTCAGATTTCAG	2440
Db	90577	ATTTCAGATAGCGATTCAGACTCAGACAGCGACTCAGACTCAGATAGCGATTCGACTCGG	90636
Qy	2441	ACAGCGATTCAGATTCAGACAGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACTG	2500
Db	90637	ATAGCGACTCAGACTCAGATAGCGATTCGCGACTCAGATAGCGATTCAGATTCGCGACTG	90696
Qy	2501	ACTTCAGATTCGCGAGTGTACTTCGGATTCAGATAGCGATTCGCGACTCAGACTCAG	2560
Db	90697	ACTTCAGATTCAGATAGCGATTCGCGACTCAGATAGCGACTCAGATTCGCGACTCAG	90756
Qy	2561	ATTTCGCGAGTGTACTCAGACTCAGACAGTGTATTTCGCGACTCAGCGAGTGTATTTCGCGATTCAG	2620
Db	90757	ATTTCGCGAGTGTATTTCAGACTCAGACAGTGTACTTCGGATTCAGATAGCGATTCGCGACTCAG	90816
Qy	2621	ATAGTGTATTCGCGACTTCGCGAGTGTACTTCGGATTCAGATAGCGACTCAGATTCGCGATAGCG	2680
Db	90817	ACAGTGTACTTCGCGACTTCGCGAGGATTCAGATTCAGACAGTGTATTTCAGACTCAGACAGTG	90876
Qy	2681	ACTTCGGATTCAGATAGCGATTCGCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAG	2740
Db	90877	ACTTCGGATTCAGATAGCGATTCGCGACTCAGATAGTGTCTCAGATAGCGACTCAG	90936
Qy	2741	AATTCAGACAGCGATTCAGATTCAGACAG-----CGACTCAG	2776
Db	90937	ACTTCAGACAGCGATTCAGATTCAGATAGTGTACTCAGATTCAGATAGCGATTCGCGACTCAG	90996
Qy	2777	ACAGTGTACTCAGATTCAGATAGTGTACTTCGGATTCAGCGAGTGTATTTCAGACTCAGGT-----	2832
Db	90997	ACAGTGTACTCAGATTCAGATAGTGTACTTCGGATTCAGCGAGTGTATTTCGAGATTCAGATTCAG	91056

```
Qy 2833 --AGTGACTCCGATTTCATCAAGTGATTCGACTCAGAAAGTGATTCAAAATAGCGATTCCG 2890
Db 91057 CGAGTGACACAGGATCAGATACGACTCAGACTCAGAAAGTGATTCCGAATAGCGATTCCG 91116
Qy 2891 AGTCAGGTTCTAACAAATATGTAGTTCGCGCTTAATTCACCTCAAAATGGTACTAATGCTT 2950
Db 91117 ATTCAAGTTCTAACAAATATGTAGTTCGCGCTTAATTCACCTCAAAAGTGGTACTAATGCTT 91176
Qy 2951 CTAATAAAATGAGGCTAAAGTAGTAAGAACCAATACCAGATACAGGTTCTGGAATG 3010
Db 91177 CAAATAAAATGAGGCTAAAGTAGTAAGAACCAATACCAGATACAGGTTCTGGAATG 91236
Qy 3011 AAGCAATACGTCACATAATTTGGGATTAATAGCAATCAATAGGTTCAATTAATCACTTTTCA 3070
Db 91237 AAGCTAATACGTCACATAATTTGGGATTAATAGCAATCAATAGGTTCAATTAATCACTTTTCA 91296
Qy 3071 GAAGAAAAAGAAAAATAAAGATAAGAAATAAGTAATAGTAATTAATTAATCAATG 3130
Db 91297 GAAGAAAAAGAAAAATAAAGATAAGAAATAAGTAATAGTAATTAATTAATCAATG 91356
Qy 3131 ATTCATGAAGAACCACTTAAGAGTGCTCTTTTACTTGGATTTTCCAAATATATGCT 3190
Db 91357 ATTCATGAAGAACCACTTAAGAGTGCTCTTTTACTTGGATTTTCCAAATATATGCT 91416
Qy 3191 TTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3250
Db 91417 TTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 91476
Qy 3251 ATAATTTGGCTTGGCGAAAAATAGGTTGAAGTAGGTTGTTAATTAGGAAAAATTAAG 3310
Db 91477 ATAATTTGGCTTGGCGAAAAATAGGTTGAAGTAGGTTGTTAATTAGGAAAAATTAAG 91536
Qy 3311 GAGAAAAATACAGTTCAAAAAATAATGCTAGTCTTTTACTTGGAGCATTATGCTATCA 3370
Db 91537 GAGAAAAATACAGTTCAAAAAATAATGCTAGTCTTTTACTTGGAGCATTATGCTATCA 91596
Qy 3371 CAAATTTGGGAAAGTAATCGTCGAGTGCAGTGGTTCTCGGGAGAAAGTAATCCATATGTA 3430
Db 91597 CAAATTTGGGAAAGTAATCGTCGAGTGCAGTGGTTCTCGGGAGAAAGTAATCCATATGTA 91656
Qy 3431 TCTGAGTCCGTTGAAAGTGAATTAATAAATAAATAATCTAGAACAGTGAAGAGTATAAG 3490
Db 91657 TCTGAGTCCGTTGAAAGTGAATTAATAAATAAATAATCTAGAACAGTGAAGAGTATAAG 91716
Qy 3491 AAAAGCTT 3498
Db 91717 GAAAAATTT 91724
RESULT 13
AX343875 LOCUS 2792 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 3 from Patent WO0198499.
ACCESSION AX343875
VERSION AX343875.1 GI:18491941
KEYWORDS Staphylococcus aureus
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
REFERENCE 1 Bacteria; Firmicutes; Bacillales; Staphylococcus.
AUTHORS Foster S.D., McDowell, P.D., Brummell, K.D. and Clarke, S.D.
TITLE Antigenic polypeptides
JOURNAL Patent: WO 0198499-A 3 27-DEC-2001;
FEATURES THE UNIVERSITY OF SHEFFIELD (GB) ; Biosynexus Incorporated (US)
Location/Qualifiers
1. .2792
/organism="Staphylococcus aureus"
/mol_type="unassigned DNA"
/db_xref="taxon:11280"
ORIGIN
Query Match 50.8%; Score 1776.4; DB 6; Length 2792;
Best Local Similarity 99.9%; Pred. No. 1.7e-257;
```

```
Matches 1788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 1 GGTACCAATAAATACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTAATTA 60
Db 1003 GGTACCAATAAATACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTAATTA 1062
Qy 61 GGTAGAGTTTCATATTAATAAATAAATAAATGTTTGCAATCAAAATCGTACGTTGCTTGA 120
Db 1063 GGTAGAGTTTCATATTAATAAATAAATAAATGTTTGCAATCAAAATCGTACGTTGCTTGA 1122
Qy 121 ATTCTTAATAATAGCAATTAATAAATAAATGTTTGTTAGTAAGTATTAATTTGTTGTAATA 180
Db 1123 ATTCTTAATAATAGCAATTAATAAATAAATGTTTGTTAGTAAGTATTAATTTGTTGTAATA 1182
Qy 181 TATCGATACAAATTAATTTGCTATTAATGCAATTTTATGTTATTAATTTCCATTAACAGAGATT 240
Db 1183 TATCGATACAAATTAATTTGCTATTAATGCAATTTTATGTTATTAATTTCCATTAACAGAGATT 1242
Qy 241 AAATATATC-TTAAAGGGTATATAGTTTAATATAAATGACTTTTTTAAAAAGAGGAAATA 299
Db 1243 AAATATATC-TTAAAGGGTATATAGTTTAATATAAATGACTTTTTTAAAAAGAGGAAATA 1302
Qy 300 AATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTGGCGTGGCTTC 359
Db 1303 AATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTGGCGTGGCTTC 1362
Qy 360 AGTGGCTTGTAGTACGTTAATCGGTTTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAG 419
Db 1363 AGTGGCTTGTAGTACGTTAATCGGTTTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAG 1422
Qy 420 TGAAAAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAAG 479
Db 1423 TGAAAAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAAG 1482
Qy 480 TAGCGTTAGTGTGCAACCTTAATAAGACAGACACAAACGTGAGTATCTAATAACATCGTC 539
Db 1483 TAGCGTTAGTGTGCAACCTTAATAAGACAGACACAAACGTGAGTATCTAATAACATCGTC 1542
Qy 540 AAACACTAATTAATGCGAAACGAGTGTGCGCAAAATCCAGCACAAACAGGAAACGACACA 599
Db 1543 AAACACTAATTAATGCGAAACGAGTGTGCGCAAAATCCAGCACAAACAGGAAACGACACA 1602
Qy 600 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 659
Db 1603 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 1662
Qy 660 GACAAAGTAATCAAGCTAATATACACCGGCAACATCAATCAAGCAATCAAAATGCGGAGGA 719
Db 1663 GACAAAGTAATCAAGCTAATATACACCGGCAACATCAATCAAGCAATCAAAATGCGGAGGA 1722
Qy 720 ATTAGTGAATCAAAACAAAGTAATGAAACGACTTTTAAATGATCTAATACAGTATCATCTGT 779
Db 1723 ATTAGTGAATCAAAACAAAGTAATGAAACGACTTTTAAATGATCTAATACAGTATCATCTGT 1782
Qy 780 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACAAACGCAAGATCTTCAAC 839
Db 1783 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACAAACGCAAGATCTTCAAC 1842
Qy 840 TGAAGCAACACCTTCAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAAGA 899
Db 1843 TGAAGCAACACCTTCAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAAGA 1902
Qy 900 TGTAGTGAATCAAGCGTTAATACAGTGGCTAGAAATGAGAGCAATTTAGTTTACGGCG 959
Db 1903 TGTAGTGAATCAAGCGTTAATACAGTGGCTAGAAATGAGAGCAATTTAGTTTACGGCG 1962
Qy 960 AGTAGTGCAGATGCAACCGGCTGGCAGATATTAAGAAATGAGTTCAGCAATGTGAC 1019
Db 1963 AGTAGTGCAGATGCAACCGGCTGGCAGATATTAAGAAATGAGTTCAGCAATGTGAC 2022
Qy 1020 AGTTGGTATTTGACTCTGGTACGCTGTGTATCGGACCAAGCAGGTTATGTCAAATGAA 1079
Db 2023 AGTTGGTATTTGACTCTGGTACGCTGTGTATCGGACCAAGCAGGTTATGTCAAATGAA 2082
```


Db 658 TAGCAAAAAGTTGAAGCAACAAATACAGATAATAAGCTTACTCAGACTTCAGACATATCGA 717
Qy 1001 ATCAGTTGACGAATGTGACAGTTGGTATTCGACTCTGGTACGACTGTGTATCCGACCAAG 1060
Db 718 ATAACTCTCAAAATGTAAACAGCACAATTTGAAGCGGAGATACGATTTACCCACATAAAG 777
Qy 1061 CAGGTTATGTCAAACTCAATATTTAGTTTTCAGTGCCTAATTTCTGCTGTTTAAAGGTGACA 1120
Db 778 CAGATATGTAAATTTAAATATTCGTTTCCAGCCCCAGATGATTTCAAGCAGGAGATA 837
Qy 1121 CATTCAAAATTAACCTGACCTAAAGAAATTAACCTTAAATGTTGTAACCTTCAACTGCTAAAG 1180
Db 838 GTATTAATAATTTACTATTCCTCAAGCAATTAACCTTAAATGTTGTAACAGCTACTGCTAAAG 897
Qy 1181 TGGCACCATAATTTGGCTGGAGATCAAGTATTTGGCAATGTTGTAATCGATAGTGTGTA 1240
Db 898 CAGCTATATTTATGGCAGGTGACCAATATTTAGCTACTGGGCAATCGATGAAGAGGAA 957
Qy 1241 ATGTTATTTATACATTTACAGACTATGTAATATCTAAAGATGATGTAAGAGCAACTTTGA 1300
Db 958 ATTAAATTTATACGTTTACAGATTATGTAACATAAATAATATTTACAGGACAGATAT 1017
Qy 1301 CCATGCCCGTTATTTGACCCCTGAAAATGTTAAAGACAGGTAAATGTGACATTTGGCTA 1360
Db 1018 CAATTCCTGGGTATTTGATCTCTAAATAATGTAACACATACGGGTAAAGTCAATTTAGAGA 1077
Qy 1361 CTGGCATAGGTAGTACAAACAGCAACAAACAGCTATTTAGTAGATTATGAAATATGTA 1420
Db 1078 CATCTATTTGGCCAAACAACTGCAAGAAACCTGTGACAGTTGATTATGAAAGTATGGAG 1137
Qy 1421 AGTTTATAACTTATCTATTAAGGTACAAATGACCAAAATCGADAAAACAAATAATACGT 1480
Db 1138 AGTTTCGTAATTTGTCAATCAAGGAATCTATTTGATCAATTTGATTAAGGTGAATACAT 1197
Qy 1481 ATCGTCAGACAAATTTATGTCAATCCAGTGAGATTAACGTTATTCGCGGTTTAAACAG 1540
Db 1198 ATAGACAAACCGTATATGTAATCTAGTAGTATGATCTGTAGTGGATCCTTATTTAAGAG 1257
Qy 1541 GTAATTTAAACCAAAATACGGATAGTATGCAATTAATAGATCAGCAAAATACAGTATTA 1600
Db 1258 GCGGTTCAATACCGGGTACCAATAGTATGTAATATAGATGACAGAAATACATCAATAA 1317
Qy 1601 AAGTATATAAGTAGATAATGACGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAA 1660
Db 1318 AAGTTTATTAAGTTAAGAAAAACCGCATCTAACAGACAGCTATTAATGTTGATCCATCAA 1377
Qy 1661 ACTTTGAGGATGTCATTAATAGTGTGAATATTAATCAATCCCAATCCAAATCAATATAAG 1720
Db 1378 ATTATGAAGATGTTACAAAGTGACGTCAAAATTTACTTTCCCTGAAAAAGGAATATATCAA 1437
Qy 1721 TAGAGTTTATACGCTGATCATCAAAATTAACACCGTATATAGTAGTTGTTAATGTC 1780
Db 1438 TCAACTTTAATCTGAAGATGATCAAAATTAATTTCCATACGTAGTTGTTAATTAATGTC 1497
Qy 1781 ATATTGATCCGAATPAGCAAAAGGTGATTTAGCTTTACGTTTCAACTTTATATGGGTATAACT 1840
Db 1498 ATGTTGATCCGAATAGTAATGGCAATTTGTTATTTACGTTTCAACATTAATGTTTACGATT 1557
Qy 1841 CGAATATAATTTGGCGCTCTATGTCATGGGACACAGAGTAGCATTTTAATTAACCGATCAG 1900
Db 1558 CTAATTTCACTTAGAGTTTCTATGGCATGGGATAATGAAGTAGAATATCATGACGGAAGTG 1617
Qy 1901 GTTCGTGTGAGCGGTATCGATAAACCAGTTGTTCTGTGAACACCTGTAGAGCTGCTGAAA 1960
Db 1618 GTAATGGTACGGGTATAGATAAACCCTGTAATACCGGATCAACCCGGGATTCGGAT---- 1673
Qy 1961 TTGAACCAATTTCCAGAGGATTCAGATTTCTGACCCAGGTTTCAGATTTCTGGCAGCGATTCTA 2020
Db 1674 -----AGTGAATGCAATTTCCGATACGACGACAGACTCAGATAGCGACGAG 1719
Qy 2021 ATTCAGATAGCGGTTTCAGATTCGGGTAGTGTATCTACATCAGATAGTGGTTTCAGATTCAG 2080
Db 1720 ACTCCGACAGCGATGCGGACTCCGATAGTGTAGTGTGATGATGATTCGACAGCGATGCACTCCG 1779

Qy 2081 CGAGTGTATTTCAGATTTCAGCAAGTGATTTCAGACTCAGCGAGTGATTTCAGATTTCAGCAAGCG 2140
Db 1780 ATAGTGACCGAGATTTCCGACAGCGATGCGAGATTCGAGATTCGAGACTCCGATAGCG 1839
Qy 2141 ATTCCGACTCAGCGAGGATTCGAGACTCAGACAAATGACTCGGATTCAGATAGCGATTCTG 2200
Db 1840 ACGCAGACTCCGATAGTGTGCGGATTCGATAGTGTGATGCGGATTCGATAGCGACGAG 1899
Qy 2201 ACTCAGACAGTGACTCAGATTCGACAGAGTACTCAGATTCAGATTCAGATAGCGATTCTGACTCAG 2260
Db 1900 ACTCAGATAGCGAGCGAGACTCCGACAGCGATGCGGACTCCGATAGTGTGCGGACTCCG 1959
Qy 2261 ACAGTGACTCAGATTCAGATAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTG 2320
Db 1960 ATAGTGATGAGATTCGACAGCGATGCGACTCAGATAGTGTGCGGAGATTCGACAGCG 2019
Qy 2321 ATTCCGACTCAGACAGCGATTCGACTCCGACAGTGTGACTCCGACTCAGACAGCGATTTCAG 2380
Db 2020 ATGCAGATTCGATAGTGTGCGGACTCCGACAGTGTGCGGACTCAGATAGTGTGCGGAG 2079
Qy 2381 ATTCCGACAGTGATTCGACTCAGATAGCGATTCGACTCAGATAGCGACTCAGATTTCAG 2440
Db 2080 ACTCCGACAGCGATGCGGACTCCGACAGCGAGATTCGATAGTGTGCGGACTCAGACTCAG 2139
Qy 2441 ACAGCGATTCAGATTCAGACAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTG 2500
Db 2140 ATAGCGATGAGATTCGATAGTGTGCGGACTCCGACAGATTCGATAGTGTGCGGACTCCG 2199
Qy 2501 ACTCAGATTCGACAGTGACTCAGATTCGAGATTCAGATAGCGATTCAGATTCGACAGTGACTCAG 2560
Db 2200 ACGCGATTCGACAGTGACTCAGATTCGATAGTGTGCGGACTCAGATAGTGTGCGGAG 2259
Qy 2561 ATTCCGACAGTGACTCAGACTCAGACAGTGTGCGGATTCAGCGAGTGATTTCGAGATTTCAG 2620
Db 2260 ATTCCGACAGCGATGCGAGATTCGACAGCGATTCGAGATTCGACAGCGAGCGAGACTCAG 2319
Qy 2621 ATAGTGATTCGAGTCCGACTCCGACAGTGACTCAGATTCGAGATTCAGATAGCGACTCAGACTCAG 2680
Db 2320 ATAGCGATGAGACTCCGACAGCGATGCGAGATTCGATAGTGTGCGGACTCCGACAGTG 2379
Qy 2681 ACTCGGATTCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGATTCAGACAGCGATTTCAG 2740
Db 2380 ATCGGACTCCGACAGTGATGCGGACTCCGATAGCGAGCGAGATTCGACAGCAATTCG 2439
Qy 2741 AATCAGACAGCGATTCAGATTCAGACAGCGACTCAGACAGTGA 2783
Db 2440 GTTCCGACAGCAATGTGAGCTCAGATAGTGTGATTCAGAAATGA 2482

Search completed: November 5, 2004, 10:04:35

Job time : 14386 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:52:30 ; Search time 2.75536 Seconds
(without alignments)
3282.472 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_425
Perfect score: 487
Sequence: 1 TLTPAYIDPENVKKTGNVT.....NVIAPVLTGNLKPNTDSNAL 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	100.0	933	2 S41539	fibrinogen-binding
2	422	86.7	989	2 D89852	fibrinogen-binding
3	113	23.2	1018	2 A32192	fibrinogen-binding
4	100.5	20.6	940	2 S19702	fibrinogen-binding
5	99.5	20.4	1092	2 T30214	fibrinogen-binding
6	88.5	18.2	1166	2 T28680	fibrinogen-binding
7	87.5	18.0	961	2 G90053	fibrinogen-binding
8	87	17.9	1141	2 E89824	fibrinogen-binding
9	84.5	17.4	1038	2 H90053	fibrinogen-binding
10	81.5	16.7	1315	2 T28679	fibrinogen-binding
11	78	16.0	202	2 G70380	probable membrane
12	76.5	15.7	1183	2 S63046	probable ribosomal
13	75.5	15.5	109	2 T38657	probable ribosomal
14	75.5	15.5	282	2 G64479	probable ribosomal
15	75.5	15.5	553	2 H64494	probable ribosomal
16	75.5	15.5	586	2 T51211	probable ribosomal
17	73.5	15.1	745	2 S13586	triacylglycerol li
18	72	14.8	3283	2 AC1018	large repetitive p
19	71.5	14.7	591	1 NUZ0F	glucose-6-phosphat
20	71.5	14.7	894	2 S54786	multifunctional be
21	71.5	14.7	1946	2 AE1449	hypothetical prote
22	71	14.6	877	2 F90070	Clumping factor B
23	70	14.4	391	2 E71125	probable aromatic-
24	70	14.4	415	2 B97085	DNA-methyltransfer
25	70	14.4	633	2 AD1116	internalin, probab
26	70	14.4	1959	2 AG1085	hypothetical prote
27	69.5	14.3	308	2 D90261	conserved hypothet
28	69.5	14.3	528	2 T41362	hypothetical prote
29	69.5	14.3	1305	2 D82923	DNA-directed RNA p

30	69	14.2	679	2 S43129	penicillin-binding
31	69	14.2	1872	2 S64851	probable membrane
32	68.5	14.1	154	2 B48230	T-cell receptor be
33	68.5	14.1	165	2 S31618	T-cell receptor be
34	68.5	14.1	995	2 S50358	hypothetical prote
35	67.5	13.9	329	2 A30760	1-phosphatidylinos
36	67.5	13.9	345	2 F86368	hypothetical prote
37	67.5	13.9	549	2 C86692	peptide-binding pr
38	67.5	13.9	1132	2 S37932	hypothetical prote
39	67.5	13.9	1324	1 VGH59	E2 glycoprotein pr
40	67.5	13.9	1385	2 D89824	hypothetical prote
41	67	13.8	174	2 AF2441	hypothetical prote
42	67	13.8	230	2 T28262	ORF MSV100 probabl
43	67	13.8	450	2 B96561	hypothetical prote
44	67	13.8	512	2 AD1132	transcription regu
45	67	13.8	533	2 D91245	malate synthase A

ALIGNMENTS

RESULT 1
S41539
fibrinogen-binding protein - Staphylococcus aureus
N;Alternate names: clumping factor
C;Species: Staphylococcus aureus
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S41539; S36630
R;McDevitt, D.; Francois, P.; Vaudaux, P.; Poster, T.J.
Mol. Microbiol. 11, 237-248, 1994
A;Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus
A;Reference number: S41539; MUID:94224142; PMID:8170386
A;Accession: S41539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-933 <WCD>
A;Cross-references: UNIPROT:Q53653; EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g39752

Query Match	100.0%;	Score 487;	DB 2;	Length 933;
Best Local Similarity	100.0%;	Pred. No. 1.2e-39;		
Matches 94;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TLTPAYIDPENVKKTGNVLTATGIGSTTANKVLVDYKYGKFNLSIKGTIDQIDKTN	60	
Db	332	TLTPAYIDPENVKKTGNVLTATGIGSTTANKVLVDYKYGKFNLSIKGTIDQIDKTN	391	
Qy	61	NTYRQIYVNPNGDVIAPVLTGNLKPNTDSNAL	94	
Db	392	NTYRQIYVNPNGDVIAPVLTGNLKPNTDSNAL	425	

RESULT 2
D89852
fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (strain
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: D89852
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Hiramatsu, K.
C;Shiba, T.; Hattori, M.; Ogawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: D89852
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-989 <KUR>
A;Cross-references: UNIPROT:Q53653; GB:BA000018; PID:gl3700678; PIDN:BA041975.1; GSPDB:G7
A;Experimental source: strain N315
C;Genetics:
A;Gene: cifa

Query Match 86.7% ; Score 422; DB 2; Length 989;

```
Best Local Similarity 84.9%; Pred. No. 2.9e-33;
Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 LTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKFKYFNLSIKGTIDQIDKTN 61
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 333 ITMPAYIDPENVTGNTVTLTGIGTNPASIKYGFHLSIKGTIDQIDKTN 392
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 62 TYRQTIYVNPISGDNVIAPLVGNLKPNTDSNAL 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 393 TYRQTIYVNPISGDNVIAPLVGNLKPNTDSNAL 425
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
A32192
fibrinectin-binding protein - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 16-Feb-1997
C:Accession: A32192
R:Signaes, C.; Raucii, G.; Joensson, K.; Lindgren, P.E.; Anantharamaiah, G.M.; Hoesek, M.
Proc. Natl. Acad. Sci. U.S.A. 86, 699-703, 1989
A:Title: Nucleotide sequence of the gene for a fibrinectin-binding protein from Staphylo
A:Reference number: A32192; MUID:89098998; PMID:2521391
A:Accession: A32192
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1018 <SIG>
A:Cross-references: GB:J04151
C:Keywords: fibrinectin binding

Query Match 23.2%; Score 11.3; DB 2; Length 1018;
Best Local Similarity 25.0%; Pred. No. 0.0059;
Matches 23; Conservative 24; Mismatches 43; Indels 2; Gaps 2;

QY 2 LTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDY-EKYGKFNLSIKGTIDQIDKTN 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 LEINLFDPKTVQNGNQTITSTNEQTSKELDVYKDGIGNY-ANLNGSIETFNKAN 359
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NTRYQTIYVNPISGDNVIAPLVGNLKPNTDSN 92
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 360 NRSFHVAFIKENGGKTSVTVTGTLMKGSNQ 391
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
S19702
fibrinectin-binding protein B - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S19702
R:Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibrinectin binding proteins in Staphylococcus aureu
A:Reference number: S19702; MUID:92111475; PMID:1837266
A:Accession: S19702
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <JOB>
A:Cross-references: UNIPROT:Q53682; EMBL:X62992; NID:G49040; PIDN:CAA44726.1; PID:G58156
C:Keywords: fibrinectin binding

Query Match 20.6%; Score 100.5; DB 2; Length 940;
Best Local Similarity 23.2%; Pred. No. 0.089;
Matches 23; Conservative 25; Mismatches 38; Indels 13; Gaps 3;

QY 2 LTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDY-----EKYGKFNLSIKGTIDQI 56
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 273 LSLNLFIDPTVTQGNQVEVKIGETVSKIFNIQYLGVRDNGW-----VTANGRIDTL 328
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 57 DKTNNTRYQTIYVNPISGDNVIAPLVGNL-----KPNTDS 91
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 329 NKVDGKFHFAYMKPNQSLSSVTVTGQVTKGNKPGVNN 367
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
```

```
T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30214
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.
Infect. Immun. 66, 2666-2673, 1998
A:Title: A fibrinogen-binding protein of Staphylococcus epidermidis.
A:Reference number: Z20781; MUID:98261511; PMID:9596732
A:Accession: T30214
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1092 <NIL>
A:Cross-references: UNIPROT:O70022; EMBL:Y17116; NID:el296734; PID:el296735; PIDN:CAA7665

Query Match 20.4%; Score 99.5; DB 2; Length 1092;
Best Local Similarity 31.5%; Pred. No. 0.13;
Matches 23; Conservative 14; Mismatches 31; Indels 5; Gaps 2;

QY 2 LTMPAYIDPENVKKTG--NVTATGIGSTTANKTVLVDYKFKYFNLSIKGTIDQIDK 58
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 389 LKLSYIDKSKVPNNYKLDVEYKTLSS--VNKTIIVEYQRPENRTANLQSMETNIDT 446
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 59 TNNTRYQTIYVNP 71
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 447 KNHTVEQTIYNP 459
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
T28680
fibrinogen-binding protein homolog - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28680
R:Joensson, E.; McCrea, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of Sta
A:Reference number: Z20510; MUID:99098700; PMID:9884231
A:Accession: T28680
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1166 <JOS>
A:Cross-references: UNIPROT:O86489; EMBL:AJ005647; NID:el318793; PID:el318794; PIDN:CAA0

C:Genetics:
A:Gene: sdrE

Query Match 18.2%; Score 88.5; DB 2; Length 1166;
Best Local Similarity 29.8%; Pred. No. 1.7;
Matches 31; Conservative 17; Mismatches 43; Indels 13; Gaps 4;

QY 2 LTMPAYIDPENV--KKTGNVTLATGIGSTTANKTVLVDYKFKYFNLSIKGTIDQIDKT 59
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 387 LTLYSYIDKKTVPNETSLNLTFTAGKETSON--VTVDYQDPWVHGDSNIQSIETKLED 444
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 60 NNTYRQTIYVNP---SGDNVIAPLV-----TGNLKPNTDSNAL 94
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 445 KQTIEQIYVNPCLKSATNTKVDIAGSOVDYGNKLGNGSTII 488
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
G90053
hypothetical protein fnbB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: G90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yurawa, H.; Kobayashi, I.; Cui, L.; Oguchi
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, R.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G90053
A>Status: preliminary
```

A;Gene: fnb

Query Match 17.4%; Score 84.5; DB 2; Length 1038;
Best Local Similarity 24.7%; Pred. No. 3-7;
Matches 21; Conservative 17; Mismatches 46; Indels 1; Gaps 1;

Qy 2 LTMPAYIDPENVKKTGNVTLATIGSTTANKTVLVDYKYGKFYNLSIKGTIDQIDTKNN 61
| : | | | : | : | : | : | : | : | : | : | : |
Db 295 LEINLFDPKTVQSNGEQKITSLKEGTEKTIPVVYPNPGVSNTYNVANGSIETFNESN 354

Qy 62 TYRQTIVNPSGDNVIAPV-LTGNL 85
| : | : | | | : | : | : | : | : | : | : | : | : |
Db 355 KFTIAYIKPMNGSNOSNTSVVTGL 379

RESULT 10
T28679
fibrinogen-binding protein homolog - Staphylococcus aureus
C;Species: Staphylococcus aureus
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28679
R;Joseffson, E.; McCrea, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.S.
Microbiology 144, 3387-3395, 1998
A;Title: Three new members of the serine-aspartate repeat protein multigene family of Staphylococcus aureus
A;Reference number: Z20510; PMID:99098700; PMID:9884231
A;Accession: T28679
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1315 <POS>
A;Cross-references: UNIPROT:O86488; EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDD:CAA04000
C;Genetics:
A;Gene: sdrD

Query Match 16.7%; Score 81.5; DB 2; Length 1315;
Best Local Similarity 31.2%; Pred. No. 9-8;
Matches 24; Conservative 12; Mismatches 34; Indels 7; Gaps 3;

Qy 7 YIDPENVKKTGN-VLTATIGISTTANKTVLVDYKYGKFYNLSIKG----TIDQIDTKNN 61
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 361 YMDADTI PVSKNDVEFNTIGNTTKTANI QPYDVVNEKNISGSAGFTETVSHGVNKEN 420

Qy 62 T--YRQTIYVPNSGDNV 76
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 421 PGYYKQTIYVPENSEL 437

RESULT 11
G70380
DNA polymerase III epsilon subunit - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: G70380
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Overby,
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; PMID:98196666; PMID:9537320
A;Accession: G70380
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-202 <AQF>
A;Cross-references: UNIPROT:C67074; GB:AE000714; NID:g2983446; PIDD:AAC07031.1; PID:g2983446
C;Genetics:
A;Gene: dnaQ

Query Match 16.0%; Score 78; DB 2; Length 202;
Best Local Similarity 31.6%; Pred. No. 2-2;
Matches 25; Conservative 10; Mismatches 32; Indels 12; Gaps 3;

Qy 1 TLTPPAYIDPENVKKTGNVTLATIGSTTANKTVLVDYEKY-----GKFYNLSIKGTI 53
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 55 TLVYFGVFPIRIKKLTGITNAMLVGOPTIEE-VLPFLPFFVGDNWGHGFVEDQIK--- 110

QY 54 DQIDKTNNTYRQTIYVNP 72
Db 111 -FINKYTKQYRGKFRNPS 128

RESULT 12
S63046
Probable membrane protein YNL106C - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N2160
C:Species: Saccharomyces cerevisiae
C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
C:Accession: S63046; S63047; S58246; S67353; S63948
R:Saiz, J.E.; Buitrago, M.J.; Soler-Mira, A.; Del Rey, F.; Revuelta, J.L.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63037
A:Accession: S63046
A:Molecule type: DNA
A:Residues: 1-1102 <SAI>
A:Cross-references: UNIPROT:P50942; EMBL:Z71382; MIPS:YNL106C
A:Experimental source: strain S288C
R:De Antoni, A.; D'Angelo, M.; Dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63047
A:Accession: S63047
A:Molecule type: DNA
A:Residues: 880-1183 <DEA>
A:Cross-references: EMBL:Z71382; MIPS:YNL106C
A:Experimental source: strain S288C
R:Saiz, J.E.; Buitrago, M.J.; Soler, A.; del Rey, F.; Revuelta, J.L.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of a 21.3 kb fragment from the left arm of yeast chromosome XI
A:Reference number: S58246
A:Accession: S58246
A:Molecule type: DNA
A:Residues: 1-1102 <SAW>
A:Cross-references: EMBL:Z50161; NID:G929846; PID:G929847
R:d'Antoni, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.
submitted to the EMBL Data Library, February 1996
A:Description: The DNA sequence of cosmid 14-13b from chromosome XIV of Saccharomyces cerevisiae
A:Reference number: S67327
A:Accession: S67353
A:Molecule type: DNA
A:Residues: 880-1183 <DAN>
A:Cross-references: EMBL:Z69382; NID:gl183941; PID:gl183968
R:Saiz, J.E.; Buitrago, M.J.; Soler-Mira, A.; del Rey, F.; Revuelta, J.L.
Yeast 12, 403-409, 1996
A:Title: The sequence of a 21.3 kb DNA fragment from the left arm of yeast chromosome XI
A:Reference number: S63948; MUID:96267765; PMID:8701612
A:Accession: S63948
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 662-887 <SAF>
A:Cross-references: EMBL:Z50161
C:Genetics:
A:Gene: SGD:INP52
A:Cross-references: SGD:S0005050; MIPS:YNL106C
A:Map position: 14L
C:Keywords: transmembrane protein
F:62-78/Domain: transmembrane #status predicted <TM1>
F:667-683/Domain: transmembrane #status predicted <TM2>

Query Match 15.7%; Score 76.5; DB 2; Length 1183;
Best Local Similarity 33.3%; Pred. No. 27;
Matches 21; Conservative 5; Mismatches 32; Indels 5; Gaps 2;

QY 12 NVKKTGNVTLATGIGSTTANKTVLDVDEYKYGK----FYNLSIKGTI 67
Db 690 NIKEVGGSTKGTGFGGITGNGKAVIRFDYGATSCFVNTHLSAGASNIDRRNDY--NNI 748

QY 68 YVN 70
Db 749 YRN 751

RESULT 13

T38657
Probable ribosomal protein, mitochondrial - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38657
R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21804
A:Accession: T38657
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-109 <MUR>
A:Cross-references: UNIPROT:Q9UT82; EMBL:AL109739; NID:e1534774; PID:CAB52271.1; GSPDB:B.04
A:Experimental source: strain 972h-; cosmid c343
C:Genetics:
A:Gene: SPDB:SPAC343.08c
A:Map position: 1
A:Genome: nuclear
A:Introns: 20/3; 104/2
C:Keywords: mitochondrion

Query Match 15.5%; Score 75.5; DB 2; Length 109;
Best Local Similarity 29.6%; Pred. No. 1.8;
Matches 29; Conservative 11; Mismatches 35; Indels 23; Gaps 5;

QY 2 LTMPAYIDPENVKKTGNVTLATGIGSTTA-----NKTVLVDYDEYKGFYNLSIKGTIDQI 56
Db 7 ITRPA-----ANATRTSPTLAMNTAKNGRAILDNKGVVVDVESMG-----DKELAKPI 56

QY 57 DKTNTNTY-----RQTIYVNPSPGDNVIAPVLTGNLKP 88
Db 57 KKLNQSIQFQHWMSMTYSPNPTVQSEIQRIID--RLPS 92

RESULT 14
G64479
hypothetical protein MJ1440 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: G64479
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Xson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: G64479
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-282 <BUL>
A:Cross-references: UNIPROT:Q58835; GB:U67585; GB:L77117; NID:gl592088; PID:AAB99451.1;
C:Genetics:
A:Map position: FOR1407290-1408138
C:Superfamily: shikimate kinase, archaeal type

Query Match 15.5%; Score 75.5; DB 2; Length 282;
Best Local Similarity 21.8%; Pred. No. 5.8;
Matches 26; Conservative 23; Mismatches 39; Indels 31; Gaps 4;

QY 2 LTMPAYIDPENVKKTGNVTLATGIGST--TANKTVLDVDEYKYGK-----FYNLSIKGTI 53
Db 73 LNYSAIVETKT-----ELPKSGLSSSSATSNAVVATFDALGEKIDELILNLGKSSF 127

QY 54 DQ-----IDKTNNTYRQTIYVNPSPGDNVIAPVLTGNLKPNTDSNAL 94
Db 128 DEKLTVTGAYDDATASYGGITITDNIETKILKDKVRDLNLVILIPNLEKNVDNRM 186

RESULT 15

H64494
hypothetical protein MJ1561 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: H64494
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: H64494
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-553 <BUL>
A:Cross-references: UNIPROT:Q58956; GB:U67596; GB:L77117; NID:G2826430; PIDN:AAB99582.1;
C:Genetics:
A:Map position: REV1538169-1536508

Query Match 15.5%; Score 75.5; DB 2; Length 553;
Best Local Similarity 36.4%; Pred. No. 13;
Matches 24; Conservative 7; Mismatches 22; Indels 13; Gaps 5;
QY 12 NVKKTGNVTATGIGSTTANKTVLVDEKY--GKFYNLSIKGTIDQIDKTNNTYRQ---- 65
Db 346 NIKDVGN--LLPGVSKTV---SFYVDVKVASAKHYKLPPIE--ISYLDTANNKYKTEKFI 398
QY 66 TIYVNP 71
Db 399 DIYVKP 404

Search completed: November 3, 2004, 19:04:08
Job time : 3.75536 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:52:30 ; Search time 15,4769 Seconds
(without alignments)
3282.472 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550
Perfect score: 2702
Sequence: 1 VGTLLIGFGLSSKEADASEN.....NGSGSGDGDKPWPEQPD 528

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2702	100.0	933	2 S41539	fibrinogen-binding
2	2466	91.3	989	2 D89852	fibrinogen-binding
3	520	19.2	961	2 G90053	hypothetical prote
4	515	19.1	940	2 S19702	fibronectin-binding
5	494.5	18.3	1166	2 T28680	fibrinogen-binding
6	494	18.3	1141	2 B89824	hypothetical prote
7	472	17.5	1092	2 T30214	fibrinogen-binding
8	467	17.3	877	2 F90070	Clumping factor B
9	452.5	16.7	1018	2 A32192	fibronectin-binding
10	441.5	16.3	1038	2 H90053	hypothetical prote
11	398	14.7	1315	2 T28679	fibrinogen-binding
12	379	14.0	1385	2 D89824	hypothetical prote
13	356	13.2	953	2 C89824	hypothetical prote
14	249	9.2	797	1 VGBEX1	glycoprotein X pre
15	248.5	9.2	1072	2 A86827	hypothetical prote
16	236	8.7	995	2 S03558	hypothetical prote
17	234.5	8.7	2271	2 F90073	hypothetical prote
18	229	8.5	750	2 T42614	probable envelope
19	228.5	8.5	3507	2 T34513	hypothetical prote
20	226.5	8.4	1777	2 T34369	hypothetical prote
21	225.5	8.3	1459	2 T45463	membrane glycoprot
22	224	8.3	1459	2 T32271	hypothetical prote
23	223.5	8.3	833	2 E90577	lipoprotein vsaC
24	223.5	8.3	967	2 S66852	hypothetical prote
25	223.5	8.3	1161	2 S57180	probable membrane
26	223	8.3	866	2 T45462	membrane glycoprot
27	223	8.3	1125	2 E90598	membrane nucleas
28	221.5	8.2	1367	1 S48478	glucan 1,4-alpha-g
29	220	8.1	1758	2 S57015	probable purine nu

30	218.5	8.1	2232	2 T34434	hypothetical prote
31	217	8.0	725	2 A41258	a-egglutinin core
32	217	8.0	792	2 S70305	hypothetical prote
33	216	8.0	796	2 T21460	hypothetical prote
34	214	7.9	1063	2 D86731	hypothetical prote
35	212	7.8	1107	2 AC0976	probable autotrans
36	209.5	7.8	971	2 B90835	probable tail fibre
37	209.5	7.8	973	2 C85693	probable membrane
38	207	7.7	528	2 I47141	gastric mucin (clo
39	207	7.7	551	2 S64314	probable membrane
40	207	7.7	1275	2 T33369	hypothetical prote
41	206.5	7.6	1229	2 T25697	hypothetical prote
42	205	7.6	888	2 T46726	secreted acid phos
43	205	7.6	1104	2 S59310	probable membrane
44	204	7.5	1169	2 S38181	floculation prote
45	203	7.5	1829	2 T24583	hypothetical prote

ALIGNMENTS

RESULT 1
S41539
fibrinogen-binding protein - Staphylococcus aureus
N;Alternate names: clumping factor
C;Species: Staphylococcus aureus
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S41539; S36630
R;McDevitt, D.; Francois, P.; Vaudaux, P.; Poster, T.J.
Mol. Microbiol. 11, 237-248, 1994
A;Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staph
A;Reference number: S41539; MUID:94224142; PMID:8170386
A;Accession: S41539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-933 <WCD>
A;Cross-references: UNIPROT:Q53653; EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g39752

Query Match 100.0%; Score 2702; DB 2; Length 933;
Best Local Similarity 100.0%; Pred. No. 4.4e-125;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VGTLLIGFGLSSKEADASENSVTQSDSASNESKNSDSSVSAAAPKTDGTTNVSDTKSSNT	60
Db	23	VGTLLIGFGLSSKEADASENSVTQSDSASNESKNSDSSVSAAAPKTDGTTNVSDTKSSNT	82
Qy	61	NGGETSVAQNPAAQOETTQSSSTNATTEETPVTGEATTTTNTQANTPATTOSSNTNABELV	120
Db	83	NGGETSVAQNPAAQOETTQSSSTNATTEETPVTGEATTTTNTQANTPATTOSSNTNABELV	142
Qy	121	NOTSNTEFTNDTNTVSSVNSPQNSTNAENVSTTQDTSTEATPSNNESAPOSTDASNKDVV	180
Db	143	NOTSNTEFTNDTNTVSSVNSPQNSTNAENVSTTQDTSTEATPSNNESAPOSTDASNKDVV	202
Qy	181	NOAVNTSAPRMRAFSLAAVAADAPAAAGTDITNLTNTVTGIDSGTTTYPHQAQYKVLNYG	240
Db	203	NOAVNTSAPRMRAFSLAAVAADAPAAAGTDITNLTNTVTGIDSGTTTYPHQAQYKVLNYG	262
Qy	241	FSVPNSAVGDTFKIITVPKELNNGVTSTAKVPPIMAGDOVLANGVIDSGNVIYFTDY	300
Db	263	FSVPNSAVGDTFKIITVPKELNNGVTSTAKVPPIMAGDOVLANGVIDSGNVIYFTDY	322
Qy	301	VNTKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKYGKFNLSIKG	360
Db	323	VNTKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKYGKFNLSIKG	382
Qy	361	TIDQIDKTNNTYQTIIYVNPSPGDNVIAVLITGNLKNPTDSNALIDQONTSIKYYKVDNAA	420
Db	383	TIDQIDKTNNTYQTIIYVNPSPGDNVIAVLITGNLKNPTDSNALIDQONTSIKYYKVDNAA	442
Qy	421	DLSESYFVNPNEDVTNSVNIITFPNPQYKVFNTPDQIITTPYIVVNGHIDPNSKGD	480
Db	443	DLSESYFVNPNEDVTNSVNIITFPNPQYKVFNTPDQIITTPYIVVNGHIDPNSKGD	502

QY 481 LALRSTLYGNSNIIWRSMWNEVAFNNGSGDGDIDKPVVPEOPDE 528
DB 503 LALRSTLYGNSNIIWRSMWNEVAFNNGSGDGDIDKPVVPEOPDE 550

RESULT 2
D89852
fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (strain
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
A:Accession: D89852
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.;
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: D89852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-989 <KUR>
A:Cross-references: UNIPROT:Q99VJ4; GB:BA000018; PID:g13700678; PIDN:BAB41975.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: clfa

Query Match 91.3%; Score 2466; DB 2; Length 989;
Best Local Similarity 91.1%; Pred. No. 1.7e-113;
Matches 481; Conservative 19; Mismatches 28; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKEADASENSVTQSDSASNESKNSDSSVSAAPKTDNTVSDTKTSNT 60
DB 23 VGTLLGFGLLSKEADASENSVTQSDSASNESKNSDSSVSAAPKTDNTVSDTKTSNT 82

QY 61 NNGETSVAQNPAQOETTQSSSTNATTEETPVTEATTTTNOANTPATQSSNTNAEELV 120
DB 83 NNGETSVAQNPAQOETTQSSSTNATTEETPVTEATTTTNOANTPATQSSNTNAEELV 142

QY 121 NOTSNETTNDNTVSSVNSPQNSNAENSVTQDTSFEAPSPNNEAPQSTDAENKDVV 180
DB 143 NOTSNETTNDNTVSSVNSPQNSNAENSVTQDTSFEAPSPNNEAPQSTDAENKDVV 202

QY 181 NOAVNTSAPRMRAPSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYVKLYNG 240
DB 203 SOAVNPSTPRMRAPSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYVKLYNG 262

QY 241 FSVNSAVKGTDFKLTVPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
DB 263 FSVNSAVKGTDFKLTVPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 322

QY 301 VNTKDDVKATLTMPAYIDPENVKITGNVTLATGISTTANKTLVDYKGYKFNLSIKG 360
DB 323 VDKNENVTANITMPAYIDPENVTGNTGNVTLTGIGTNTASKTLVDYKGYKQFHNLSIKG 382

QY 361 TIDQIDKTNNTYRQTIYVNSPQNSVIAPLVTCNLKPNNTDSNALIDQOANTSIKVYKVDNAA 420
DB 383 TIDQIDKTNNTYRQTIYVNSPQNSVIAPLVTCNLKPNNTDSNALIDKNTDVKYVDNAN 442

QY 421 DLSSEYFVNPNFEDVNSVNIITPNPNQYKVEFNTPDQITTPYIVVNGHIDPNSKGD 480
DB 443 DLSSEYFVNPNFEDVNTNQVRIISFPNANQYKVEFTDDDDQITTPYIVVNGHIDPASTGD 502

QY 481 LALRSTLYGNSNIIWRSMWNEVAFNNGSGDGDIDKPVVPEOPDE 528
DB 503 LALRSTLYGNSNIIWRSMWNEVAFNNGSGDGDIDKPVVPEOPDE 550

RESULT 3
G90053
hypothetical protein fnbB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: G90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.;
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: G90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-961 <KUR>
A:Cross-references: UNIPROT:Q99RD3; GB:BA000018; PID:g13702452; PIDN:BAB43593.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnbB

Query Match 19.2%; Score 520; DB 2; Length 961;
Best Local Similarity 26.8%; Pred. No. 3.9e-18;
Matches 149; Conservative 89; Mismatches 201; Indels 118; Gaps 15;

QY 1 VGTLLGFGLLSKEADASENSVTQSDSASNESKNSDSSVSAAPKTDNTVSDTKTS 57
DB 21 LGTIVVGMGQEKAAASEQNNTTVESGSSATSKA-----SETQIT 63

QY 58 SNTNNGETSVAQNPAAQOETTQSSSTNATTEETPVTEATTTTNOANTPATQSSNTNAE 117
DB 64 TNNVNTIDETQSYSATSTEQPSKSTQVTEEAPTVQAPKVETEMKS----- 110

QY 118 ELVNQTSNETTNDNTVSSVNSPQNSNAENSVTQDTSFEAPSPNNEAPQSTDA 174
DB 111 -----QEDLPSEKADKETTGTQVDIAQPSN----- 136

QY 175 SNKDVVNAVNTSAPRMR-----AFSLAAVAADAPAGTDTITNQLTNVTGID---SG 224
DB 137 -----VSEIKPRMKRSADVTAVSEKEVAEAKATGTDVTKNKEVTESSLEGHNKDS 187

QY 225 TTVPVPHQAGYVKLYNGFSVPNSAVKGTDFKLTVPKELNLTGVTSTAKVPPIMAG--DQVL 282
DB 188 NVNPNHNAQRTLVYKWKFGGKAGDYDFTLSDNVETHGISTLRKYPEIKSSDEKVM 247

QY 283 ANGVIDSQNVITYFTDYVNTKDDVKATLTMPAYIDPENVKITGNVTLATGIGSTANKT 342
DB 248 ANGVINERTTRYFTDYINNKDULTAELNLNLFIDPTTVTKQGSQKVEVTLGQNKVSK 307

QY 343 VLVDY-----EKYKGFYNLSIKGTIDQIDKTNNTYRQTIYVNSPQNSVIAPLVTCNLKEN 397
DB 308 FDIKYLGVKDRMG-----VTVNGRIDTLNKEEGKFSHPAYVYKPNQSSSTSVTVTGQVTS 363

QY 398 TDSNALIDQOANTSIKVYKVDNAAIDSESYFY--NPENFEDVNSVNIITFPNPNQYKVEF 454
DB 364 YKQSA---NNPTVKVYKHIGSDLAESVYAKLDDTSKFEDVTEKVNLSYTSNGGYTLNL 419

QY 455 NTPDDQITTPYIVVNGHIDPNSKGDALRSLTLGYNSNIIR-----SMWDNEVA 506
DB 420 GDLDN--SKDVIKYEGEYDQNAK-DLNFRTLHLSGYHKYYPYYPYYPYVQLTWNNGVA 476

QY 507 FNNGSGSDGIDKPVVP 523
DB 477 FYSNAKGGDKKPNDD 493

RESULT 4
S19702
fibrinectin-binding protein B - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
A:Accession: S19702
R:Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibrinectin binding proteins in Staphylococcus aureus
A:Reference number: S19702; PMID:92111475; PMID:1837266
A:Accession: S19702
A:Status: preliminary


```
QY 61 NNGETSAQNPAQOQTSSSTNAATTEPTVTEATTTTNTQANTPATTOSSNTNAE-EL 119
Db 85 ENDST---NPKKETNTDSOPEAKEEST---TSSTQOQONNVATTEKPKQNIKEN 135
QY 120 VNQTSNETTNDNTV-----SSVSPQNSTNAENVST---T 152
Db 136 VKPSTDKTATETSTVILEBEKKAPNTNNDVTTKPSTBSIQTKPTTPOBSTNIENSQOQPT 195
QY 153 TQDSTESTAPNSESAP-----QSTDASNKDVVQAVNTSAPR 190
Db 196 PSKVDNQVTDATNPKPEPVNSKEBELKNPEKLKELVRNNDNTDRKRVATPSTVAPKR 255
QY 191 MRAFSLAAVAADAPAAAGTDIINQTLT---NVTVGIDSGTIVYPHQAGVYKLVNFGFVSPTS 246
Db 256 LNAKMRFAVAQAAVASNNVNDLITVTQTIKVGDGKNVAAAHDKDIEYDTEITDNK 315
QY 247 AVKGDTFKI-----TVPKELNLNGVTSTAKVPPIMAGD---QVLANGVID--SDGNVIYTF 297
Db 316 VKKGDWTWTINYNKVNIPSDL-----TDKNPDIDITDPSGEVIKAGTFDKRATKQITVTF 368
QY 298 TDYVNTKDDVKATLMPAYIDPENV--KKTGNVTLATGIGSTTANKTVLVYDEYKGYKFN 355
Db 369 TDYVDKYEDIKARLVLYIDKQAVNETSNLTLFATAGKETSON--VSDYIQDPMWIGD 426
QY 356 LSIKGTIDQIDKNTNTRQTIYNVP---SGDNVIAPVL-----TGNLKPNTDSNALIDQ 406
Db 427 SNIQSIFTKLDENKQITBQIYVNPPLKKTATNTKVDIAGSQVDDYGNIKLNGS--TIID- 484
QY 407 QNTSIKVKVUNAADLSESYFV--NPENFEDVYNSVNTFPNPQVKBVFNTPD--DQITTP 464
Db 485 QNTEIKVKVNPQOQLPGSNRIYDFSQVEDYTSQ-----FDNKKGSNNVATLDFGDINSA 540
QY 465 YIVVVGHHIDNSKGLD-----ALRST-LVGYNSNIIWRSMGMDNEVAFNNGSGSGDGI 517
Db 541 YIIKVKSYTSDGELDIAOGTSMRTIKIGY-----YNYAGYSNFIVTNDTGGGDT 595
QY 518 DKP 520
Db 596 VKP 598
```

RESULT 7

```
T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C/Accession: T30214
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.
Infect. Immun. 66, 2666-2673, 1998
A>Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.
A/Reference number: 220781; MUID:98261511; PMID:9596732
A/Accession: T30214
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1092 <NIL>
A/Cross-references: UNIPROT:Q70022; EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA766
```

```
Query Match 17.5%; Score 472; DB 2; Length 1092;
Best Local Similarity 26.5%; Pred. No. 1e-15;
Matches 157; Conservative 103; Mismatches 238; Indels 94; Gaps 25;

QY 1 VGTLLIGFLLSSKREADASENSV-----TQSDSASNEKSNDSVSAAPKTDD 48
Db 36 IGATLLFG-LGHNEAKAEENSVDQVKDSNTDDELSNDSQSDSEKNDVNNQSNINTDD 94
QY 49 TNV-----SDTKSSNTN---NGETSAQNP-----AQOETQSSSTN 83
Db 95 NNQIIEETNNYDGIKRSRDETSTNVNDENEATFLQKTPQDNTLITEBEVKESVSE 154
QY 84 ATTEETPTVTEATTTTNTQANTPATTOQS--SNTNAEELVN---QTSNETTNDNTV--- 135
Db 155 SSNSIDTAQOPSHTTINRESVQTSNDVSDHVSDFANSIKESNTSGKEENTIEQPN 214
```

```
QY 136 -----SSVSPQNSTNAENVSTTQDSTTEATPSN---NESAPQSTDSAGNKDVVQAVNTS 187
Db 215 KVKESTTSQSGYSTYNIDEKISNQDELLN-LPINEYENKARPLSTTSQAQSIKRVTVN- 271
QY 188 APRMRARSLAAVAADAPAAAGTDIINQTLTNTVTVGI--DSGTTVYPHQAGVYKLVNFGFVSPTS 246
Db 272 -----QLAAEQSGSNVNHILKIVTDQ--SITEGVDDSEGVIKAHDAENLIYDVTFEVDDK 322
QY 247 AVKGDTFKIYTPKELNLNGVTSTAKVPPIMAGD--QVLANGVIDS--DGNVIYTEFDYVNTK 304
Db 323 VKSGDWTMTVDIDKNTVPSDLTSTFIPKIKNSGEIITAGTYDNKNKQITFTTIDYVDKY 382
QY 305 DDVKATLTMPIAYIDPENVKKTG--NVTLATGIGSTTANKTVLVYDEYKGYKFNLSIKGT 361
Db 383 ENIKAKHLKLTSTVIDKSKVPNNNTKLDVEYKLTALSS--VNKTIITVEYQRPNRNRTANLQSM 440
QY 362 IDQIDKNTNTRQTIYNVPSGDNVIAPVL--TGNLKPNTDSNALIDQOQNTSIKVKYKVDNAAD 421
Db 441 FTNIDTKNHTVEQTIYINPL--RYSAKETNVNISGNGDEGSTIIDDSTIIEKVKYKVGDNQN 498
QY 422 LSESYFV--NPENFEDVYNSVNTFPNPQVKBVFNTPDQITTPYIVVVGHHIDPNSKGD 480
Db 499 LPDSNRIYDSEYEDVINDYQAQLGNNDVNVNIFG----NIDSPYIIKVKISKYDNP--KDD 553
QY 481 -----LALRSTLYGYNSNIIWRSMGMDNEVAFNNGSGSGDGIKFPVVPQ 525
Db 554 YTTIQQVTVMQTINEYTG--FRTASVDNTIAFSTSSGQGG--DLP--PEK 600
```

RESULT 8

```
F90070
Clumping factor B [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: F90070
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: F90070
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-877 <KUR>
A/Cross-references: UNIPROT:Q99R07; GB:BA000018; PID:gl3702588; PIDN:BAB43728.1; GSPDB:G6
A/Experimental source: strain N315
C/Genetics:
A/Gene: clfB
```

```
Query Match 17.3%; Score 467; DB 2; Length 877;
Best Local Similarity 25.8%; Pred. No. 1.4e-15;
Matches 150; Conservative 109; Mismatches 204; Indels 118; Gaps 27;

QY 1 VGTLLIGFLLSSKREADASE--NSVTQSDASNEKSNDSVSAAPKTDDTNNVSDTKTSS 58
Db 29 VGATILFG-IGNHOQAQASEQSNDDTQS--SKNSADSEKNNMIETPQLTNTANDTSDISA 86
QY 59 NTNNGETSAQNPAQOQTQSSSTN--ATTEETPTVTEATTTTNTQANTPATTOSSNTNAE 117
Db 87 NTNSANVDSTTKPMSTQTSNTTITTEPATSTNETP-----OPTAKNQAT--AAKMQDQVTPQ 140
QY 118 ELVNQTSNETTNDNTNVTSSVNSPQNSTNAENVSTTQDSTTEATPSNSESAPQSDASNK 177
Db 141 EANSQVDNKT--NDANSI-----ATNSE-----LKNSQLDLDP---QSSPQT----- 178
QY 178 DVVNQAVNTSAPRWRAPSLAAVAADAP-----AGDTITNQLNTVTVGIDSGTIVPHQ 231
Db 179 --ISNAOQTSKPSVRTRAVRSLAAVEFPVVAADAKGTNVNDKVTASNFLEK--TTFDPNQ 235
QY 232 AGYVKLVNFGFVSVPNSAVKGDTFKLTVPKELNLNGVTSTAKVPPIMAGDOVLANGVID--- 288
Db 236 SGNTFMAANFTVTDKVKSGDYF-----TAKLP-----DSLGNQGVVDYSN 275
```


A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-953 <KUR>
A:Cross-references: UNIPROT:Q99W48; GB:BA000018; PID:BA01750.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: sdrC

Query Match 13.2%; Score 356; DB 2; Length 953;
Best Local Similarity 25.8%; Pred. No. 4.2e-10;
Matches 140; Conservative 81; Mismatches 217; Indels 104; Gaps 23;

QY 1 VGTILGFLLSKEADASENSVTQSDASNEKSDSSVSAAAPKTDITNVSDFTKSSNT 60
DB 35 VGTTLIFG-LSGHEAKAAETHGELNLSKNETTA--PSENKTKTEKVDLSRQLK----- 84
QY 61 NNGETSVACNPAQOETTQSSSTNATTEETPVTEGATTTTNCANTPATQSSNTNABELV 120
DB 85 -NTQTATADQP---KVTMGDS--ATVKE-----TSSNMOSP-----QNATASQT 123
QY 121 NQTSNETTFND--TNTVSSVNSPQNSIABNVSTTQDTSTEATPSNNEAPQSDASNKD 178
DB 124 TQTSNVTT-NDKSSITTSYNETDKSNLTQAKNVST-----PKTTTIKQRA 167
QY 179 VVQAVNTSAPRMRAFLAALAAADAAPAGTHDITNQL--TNTVTGIDSG-----TTVY 228
DB 168 LNRMAVNTVA-----APOOGTNVNDKVHFTNIDIAIDKGVNKTTCNTGEFW 213
QY 229 PHQAGYVKLNYGFSVENSACKDTEKITVPKELNLNGVTSTAKVPPIM-AGDOVLANGVI 287
DB 214 ATSSDVLKUKANYTTIDDSVKEGDTFTPKYQYFRPGSVRLPSQTONLYNAQCNIIAGIY 273
QY 288 DSDGN-VIYTFDVTNKKDDVKATLTMPAYIDPEN--VKKTG---NVTLATGIGSTTANK 341
DB 274 DSKTNTTITFTNYVDQYTNVSGSFEQVAFKAGENATTDKAYKMEVTL-----GNDITYSK 329
QY 342 TVLDYKYGKFNYSIKGTIDOIDKTNNTYROTIIYNPSGDNIAPVLTGNL----KNT 398
DB 330 DIVIDYGNKGQQLISSTNYINNEDLSRN--MTVYVYVQPKKTYTKETFTVNLITGYKENP 386
QY 399 DSNALIDQNTSIKVKYKVNAAADLSESYFVNPNFEDVNTSVNITFFPNQ-YKVEFNTP 457
DB 387 DAK-----NPKIYEVTDQNFVDSFTPTSKLKVDTGQDFVIYSNDNKATVDDLNG 438
QY 458 DQIITPTPYVWNGHGDPSKGLALRSLTGLYNSNIWRSMSWDNEVAFNNGSGSGDI 517
DB 439 QSSSDKQYIIQVAYPDNSSTDNGKIDYTLTQNG-----KSSWSNSVNVNGSSTANGD 493
QY 518 DK 519
DB 494 QK 495

RESULT 14
VGBEX1
glycoprotein X precursor - equine herpesvirus 1 (strain Ab4p)
C:Species: equine herpesvirus 1
A:Note: host Equus caballus (domestic horse)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C:Accession: H36802
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
A:Submitted to GenBank, March 1992
A:Description: The DNA sequence of equine herpesvirus-1.
A:Reference number: A36805
A:Accession: H36802
A:Molecule type: DNA
A:Residues: 1-797 <TEL>
A:Cross-references: UNIPROT:P28968; GB:M66664; NID:G330791; PID:AA02506.1; PID:G330862
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A:Title: The DNA sequence of equine herpesvirus-1.
A:Reference number: A41831; MUID:92295566; PMID:1318606
A:Contents: annotation; possible protein-coding frames

A>Note: neither amino acid nor nucleotide sequence is given
C:Genetics: 71

C:Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein homolog
C:Keywords: glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-797/Product: glycoprotein X #status predicted <MAT>
F:23-465/Region: serine/threonine-rich
F:489-797/Domain: equine herpesvirus 1 glycoprotein homology <EHG>
F:766-790/Domain: transmembrane #status predicted <TMN>
F:590/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 9.2%; Score 249; DB 1; Length 797;
Best Local Similarity 21.8%; Pred. No. 5.8e-05;
Matches 114; Conservative 80; Mismatches 238; Indels 90; Gaps 17;

QY 3 TLIGFLLSSKEADASENSVTQSDASAKESKNDSSS-----VSAAPKTDDTNNVDTK 55
DB :
16 SIYAIGSTTTTETTTSSSTSGSGOSTSGTNNSSSPPTSPPTSSSPPTGHTSSPSS 75
QY 56 TSSNTNGETSVAQNPAQQETTQSSTNAATTBETPWGEATTT-----TNQANTPA 107
DB :
76 TSTQSSSTAATSSASPSTASSSTSTPTGSTETTTTTPTASTTTPTTTAAPTAAATTA 135
QY 108 TTQSSNTNAE---ELVNQTSNETTFDNTVSSVNSPQNSTNAENVSTTQDTSTEAAPS 164
DB :
136 VTAAASTSAETTTATATATSTPTTTTPTSTTTTATTTVPPTA---STTTDTTAAATTA 192
QY 165 NESAPQSTDASKDVNQAVNTSAPRMRAFSLAAVAADAPAAGTD-----ITNQLIN 216
DB :
193 ATTTAAATTTAA---TTTAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 249
QY 217 VTVGIDSGTGVVPHOAGVVKLYGFVSNSAVKGDTFKITVPKELNLNGVTSTAKVPPIM 276
DB :
250 TAATTTAATTTGSPSTGSGSTT---CASTSPSATATSATSTSTSTSAATSTSTPTPS 306
QY 277 AGDQVLANGVIDSGNVIIYFTDYVNTKDDVKATLTPAYIDPENVKKTGNVTLAFGIS 336
DB :
307 AATS--AEStTEAP-----TGTPPTDTTTPSEAT--TATTSPESTTVSASTTSAITTAF 356
QY 337 TTANKTVLVDEYKGYKNLSIKGTIDQDKNNYR--QTIVNPSGDNVIAPVLTG-- 392
DB :
357 TTESH-----SPDSSTGSGTSTAEPSSFTLTLPSTATPSTDOQTGSS 398
QY 393 -----NLKPMTDSNALIDQONTSIKVYKDNAADLSSEYFNPNFE--DVTNSV 440
DB :
399 ASTESDSDSTSVPTTGESTTESSTT-----EASNLGSTVESTALETPDGNTIS 452
QY 441 NITFPNPQYKVEF----NTPDDQITPTPIVVVNGHIDPNSK 478
DB :
453 GNTTSPSPRTSPADTQOTPDNGVSTQH-TTINDHTTANAQ 493

RESULT 15
A86827
hypothetical protein yqfG [imported] - Lactococcus lactis subspecies lactis (strain IL1403)
C:Species: Lactococcus lactis subspecies lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86827
R:Bolotin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp R
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: A86827
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1072 <STO>
A:Cross-references: UNIPROT_Q9CF64; GB:AE005176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GB
A:Experimental source: strain IL1403
C:Genetics: yqfG
A:Gene: yqfG

Query Match 9.2%; Score 248.5; DB 2; Length 1072;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:52:30 ; Search time 27.3484 Seconds
(without alignments)
3282.472 Million cell updates/sec

Title: US-09-679-643-2

Perfect score: 4725

Sequence: 1 MNMKKEKHAIRKKSIGVAS.....LASIGSILLFRKKENKDKK 933

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4725	100.0	933	2 S41539	fibrinogen-binding
2	4399	93.1	989	2 D89852	fibrinogen-binding
3	1638.5	34.7	877	2 F90070	Clumping factor B
4	1628	34.5	1092	2 T30214	fibrinogen-binding
5	1414	29.9	1166	2 T28680	fibrinogen-binding
6	1395.5	29.5	1385	2 D89824	hypothetical prote
7	1354.5	28.7	953	2 C89824	hypothetical prote
8	1320.5	27.9	1141	2 E89824	hypothetical prote
9	1110.5	23.5	1315	2 T28679	fibrinogen-binding
10	963	20.4	903	2 AG1299	probable peptidogl
11	957.5	20.3	882	2 AG1671	probable peptidogl
12	799	16.9	2271	2 F90073	hypothetical prote
13	688.5	14.6	961	2 G90053	hypothetical prote
14	682	14.0	940	2 S19702	fibrinectin-bindin
15	659.5	14.0	2570	2 T17451	fimbriae-associate
16	635.5	13.4	406	2 S38170	SRP40 protein - ye
17	614	13.0	1018	2 A32192	fibrinectin-bindin
18	611	12.9	1038	2 H90053	hypothetical prote
19	597.5	12.6	640	2 A54502	S antigen precurs
20	577	12.2	4776	2 E95206	cell wall surface
21	569.5	12.1	1063	2 D86731	hypothetical prote
22	500	10.6	334	2 A54138	acidic repetitive
23	494	10.5	1428	2 T08852	lustrin A - Califo
24	487.5	10.3	3394	2 T18501	hypothetical prote
25	479.5	10.1	1664	2 T18262	S-layer protein -
26	461	9.8	1217	2 S52714	sericinB - silkw
27	445	9.4	888	2 T46726	secreted acid phos
28	430.5	9.1	1459	2 T32271	hypothetical prote
29	426.5	9.0	833	2 E90577	lipoprotein vsaC [

30	420	8.9	792	2 T42963	hypothetical prote
31	407	8.6	2232	2 T34434	hypothetical prote
32	405.5	8.6	937	2 S58125	hypally regulated
33	401	8.5	2910	2 T28156	DNA-directed RNA p
34	388	8.2	286	2 C61615	sericin MG-2 - gre
35	379	8.0	1072	2 A86827	hypothetical prote
36	378	8.0	797	2 A36811	hypothetical prote
37	377.5	8.0	2481	2 D90011	FntB protein (impo
38	377	8.0	1192	2 A71623	probable secreted
39	374	7.9	1419	2 T30531	agglutinin-like ad
40	374	7.9	1829	2 T24563	hypothetical prote
41	373.5	7.9	670	2 T28391	ORF MSV230 hypothe
42	372.5	7.9	695	2 S27390	calcium-binding pr
43	368.5	7.8	955	2 T18435	hypothetical prote
44	363	7.7	489	2 A45988	dentin matrix acid
45	363	7.7	685	2 E71609	hypothetical prote

ALIGNMENTS

RESULT 1

S41539

fibrinogen-binding protein - Staphylococcus aureus

N;Alternate names: clumping factor

C;Species: Staphylococcus aureus

C;Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S41539, 836630

R;McDewitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.

Mol. Microbiol. 11, 237-248, 1994

A;Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus

A;Reference number: S41539; MUID: 94224142; PMID: 8170386

A;Accession: S41539

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-933 <MCD>

A;Cross-references: UNIPROT: Q53653; EMBL: Z18852; NID: g397525; PID: g39752

Query Match 100.0%; Score 4725; DB 2; Length 933;

Best Local Similarity 100.0%; Pred. No. 1.4e-202;

Matches 933; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNMKKEKHAIRKKSIGVASVLVTGLIGFLLSKKEADASENSVTQSDSASNEKSDSS	60
DB	1	MNMKKEKHAIRKKSIGVASVLVTGLIGFLLSKKEADASENSVTQSDSASNEKSDSS	60
QY	61	SVSAAPKTDITVSDTKTSNTNGETSV AQNPAQQTQSSSTNATETPTVGEATTT	120
DB	61	SVSAAPKTDITVSDTKTSNTNGETSV AQNPAQQTQSSSTNATETPTVGEATTT	120
QY	121	TTNQANTPATQSSNTNABELVNCSTNETTFDNTVSSVNSPQNSTNAENVSTQDTST	180
DB	121	TTNQANTPATQSSNTNABELVNCSTNETTFDNTVSSVNSPQNSTNAENVSTQDTST	180
QY	181	EATPSNNEAPQSTDAASNDVQAVNTSAPRMRASLA AAVADAPAAAGTDITNQLTNVT	240
DB	181	EATPSNNEAPQSTDAASNDVQAVNTSAPRMRASLA AAVADAPAAAGTDITNQLTNVT	240
QY	241	VGIDSGTGVYVPHQAGVYKLYGFSVPNSAVKGTDFKFTVPEKELNGLVSTAKVPPIMAG	300
DB	241	VGIDSGTGVYVPHQAGVYKLYGFSVPNSAVKGTDFKFTVPEKELNGLVSTAKVPPIMAG	300
QY	301	DOVLANGVIDSDGNVLYFTDYNTKDDVKATLTMPAYIDPENVKKTGNVLTATGISTT	360
DB	301	DOVLANGVIDSDGNVLYFTDYNTKDDVKATLTMPAYIDPENVKKTGNVLTATGISTT	360
QY	361	ANKTVLVDYKYGKFNLSIKGTDIDKTNNTYRTIYVNPQSDNVIAPLVGLNKLKPT	420
DB	361	ANKTVLVDYKYGKFNLSIKGTDIDKTNNTYRTIYVNPQSDNVIAPLVGLNKLKPT	420
QY	421	DSNALIDQQNTSIKVKVDNAADLSSEYFVNPFEDVTNSVNIITFPNPQYKVEFTPD	480
DB	421	DSNALIDQQNTSIKVKVDNAADLSSEYFVNPFEDVTNSVNIITFPNPQYKVEFTPD	480

```
QY 481 DOITTPYIVVNGHIDPNSKGDALRSTLYGNSNIIWRSMWNEVAFNNGSGSGDGID 540
Db 481 DOITTPYIVVNGHIDPNSKGDALRSTLYGNSNIIWRSMWNEVAFNNGSGSGDGID 540
QY 541 KPVPPEQDEPGEIEPIPEDSDSPGSDGSDNSDSDSGSDGSDSTSDSGSDASDSASA 600
Db 541 KPVPPEQDEPGEIEPIPEDSDSPGSDGSDNSDSDSGSDGSDSTSDSGSDASDSASA 600
QY 601 SDSASASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 660
Db 601 SDSASASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 660
QY 661 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 720
Db 661 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 720
QY 721 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 780
Db 721 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 780
QY 781 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 840
Db 781 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 840
QY 841 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 900
Db 841 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 900
QY 901 SEDEANTSLIWGLLASIGSLLLFRKKENKDKK 933
Db 901 SEDEANTSLIWGLLASIGSLLLFRKKENKDKK 933

RESULT 2
D89852
fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (strain
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D89852
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: D89852
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-989 <KUR>
A:Cross-references: UNIPROT:Q99WJ4; GB:BA0000018; PID:g13700678; PIDN:BABA1975.1; GSPDB:G
A:Experimental source: strain N315
A:Genetics:
A:Gene: c1fa

Query Match 93.1%; Score 4399; DB 2; Length 989;
Best Local Similarity 88.6%; Pred. No. 4.6e-188;
Matches 876; Conservative 27; Mismatches 30; Indels 56; Gaps 2;
QY 1 MNMKEKEHAIKKIGVASVLVGLTIGFLLSSKEADASENSVTQSDASNESKSDSS 60
Db 1 MNMKEKEHAIKKIGVASVLVGLTIGFLLSSKEADASENSVTQSDASNESKSDSS 60
QY 61 SVSAAPKTDITVSDTKTSNTNNGETSVAQPAQETTQSSSTNATTEETPVTEATT 120
Db 61 SVSAAPKTDITVSDTKTSNTNNGETSVAQPAQETTQSSSTNATTEETPVTEATT 120
QY 121 TTNQANTPATTQSSNTNABELVNTSETTTFNFTNTVSVNSPQNSNAENVSTTQDST 180
Db 121 TTNQANTPATTQSSNTNABELVNTSETTTFNFTNTVSVNSPQNSNAENVSTTQDST 180
QY 181 EATPNNESAPQSTASNKDVVNOAVNTSAPRMRAFSLAAVAADAPAAAGTDTITNQLTNYT 240
Db 181 EATPNNESAPQSTASNKDVVNOAVNTSAPRMRAFSLAAVAADAPAAAGTDTITNQLTNYT 240
```

```
Db 181 EATPNNESAPQNTDASNKDVVQAVNPFSTPRMRAFSLAAVAADAPAAAGTDTITNQLTNYT 240
QY 241 VGIDSGTTVYPHQAGYVKLYGFSVPNSAVAGDTEFKITVPKELNLNGVTSSTAKVPPIMAG 300
Db 241 VTIDSGTTVYPHQAGYVKLYGFSVPNSAVAGDTEFKITVPKELNLNGVTSSTAKVPPIMAG 300
QY 301 DQVLANGVIDSGNVIYFTTDDVVKATLTMPAYIDPENVKKTGNVLTATGISTT 360
Db 301 DQVLANGVIDSGNVIYFTTDDVVKATLTMPAYIDPENVKKTGNVLTATGISTT 360
QY 361 ANKTVLVYKFKYKYNLSIKGTIDQIDKTNNTYQTIYVNPNGNVIAPVLTGNLKPT 420
Db 361 ASKTVLIDYKYGQPHNLSIKGTIDQIDKTNNTYQTIYVNPNGNVIAPVLTGNLKPT 420
QY 421 DSNALIDQNTSISKYVKVDNAADLSESYFVNPFENFEDVTNSVNIITFPNPNQYKVFENPTD 480
Db 421 KSNALIDAKNTDIDKVRVDNANDLSESYFVNPFENFEDVTNQVRIISFPKANQYKVFETDD 480
QY 481 DQITTPYIVVNGHIDPNSKGDALRSTLYGNSNIIWRSMWNEVAFNNGSGSGDGID 540
Db 481 DQITTPYIVVNGHIDPASTGDLALRSTFYGSDNFIRWSMWDNEVAFNNGSGSGDGID 540
QY 541 KPVPPEQDEPGEIEPIPEDSDSPGSDGSDNSDSDSGSDGSDSTSDSGSDASDSASA 600
Db 541 KPVPPEQDEPGEIEPIPEDSDSPGSDGSDNSDSDSGSDGSDSTSDSGSDASDSASA 600
QY 601 SDSASASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 622
Db 601 SDSASASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 660
QY 623 -----SDNDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 676
Db 661 SDSASASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 720
QY 677 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 736
Db 721 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 780
QY 737 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 796
Db 781 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 840
QY 797 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 853
Db 841 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 900
QY 854 -----DSESASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 904
Db 901 STSDTSGSDNDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
QY 905 ANTSLLIWGLLASIGSLLLFRKKENKDKK 933
Db 961 ANTSLLIWGLLASIGSLLLFRKKENKDKK 989
```

```
RESULT 3
F90070
Clumping factor B [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F90070
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: F90070
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <KUR>
A:Cross-references: UNIPROT:Q99R07; GB:BA0000018; PID:g13702588; PIDN:BABA3728.1; GSPDB:G
A:Experimental source: strain N315
```

C;Genetics:
A;Gene: clfb

```
Query Match      34.7%; Score 1638.5; DB 2; Length 877;
Best Local Similarity 40.4%; Pred. No. 1.3e-65;
Matches 398; Conservative 150; Mismatches 262; Indels 175; Gaps 33;

QY 5 KKEKHAIRKKSIGVASVLVGTGILGGLSSKEADASE--NSVTQSDASNESKNSDSSV 62
DB 11 KQKISIRRTVGTTSVIGATILFG-IGNHQQAQSEQSDNTTQS-SKNWASADSEKNM 68
QY 63 SAAPKTDITNVSDTKSSNTNNGETSVAQNPAAQOETTQSSSTN-ATTEPTPTVGEATTTT 121
DB 69 IETPQLNTTANDSDISANTNSANVDSTTKPMSQTSTNTTTPASINETP-----QPTAI 124
QY 122 TNOANTPATQSSNTNAELVNQTSNETTENDTNTVSSVNSPQNSNAENVSTQDTSTE 181
DB 125 KNOAT--AAKMQDQTVFPQEAANSVDNKT--NDANSI-----ATNSE-----LKNSQTL 169
QY 182 ATPSNNEAPQSDASNKDVVNOAVNTSAPRMRAFSLAAVAADAP-----AAGTDTITNQ 235
DB 170 DLP--QSSPQT-----ISNAQGTSKSVTRAVRSLAVAEPPVNAADAKGTNVNDK 218
QY 236 LTNVTVGIDSGTIVYHQAGYVKLNYGFSVPSAVKGDTFKITVPKELNLNGVTSTAKVP 295
DB 219 VTASNFKLEK-TTFDPNQSNGTMAANFTVTDVKYSGDYF-----TAKLP 262
QY 296 PIMAGQVLANGVID-----SDGNVI-----YTFDDYVNTKD 327
DB 263 -----DSLTKNGVDVYSSNNTMPIADIKSTNGDVAKATYDILTKYTFVTFDYVNKE 317
QY 328 DVKATITMPAYIDPENVKKTGNVTLATGIGSTTANKTVLDVDEK---YGFKNLSIKGT 383
DB 318 NINGQSLPLFTBRAPKPSGTYDANINADIEMFNKITVYSPAGIDKPNGANISSQ 377
QY 384 IDQIDKTN--NTRYQTIYVNS-----GDNVIAPLVTGLNKLKNTDSNALIDQNTSIKYYK 437
DB 378 IIGVDTASGQNTYKQTVFVNPQKQVLNLTWV--YIKGYQDKIESSGKVSATDTKLRIE 435
QY 438 VDMAADLSESYFNP--ENFEDVTNSV--NITFPNPQYKVFENTPDDQITTYIIVVNG 493
DB 436 VNTSKLSLSDYYADPNDSNLKEVTDQPKRIYYEHPNVAIKFG-----DITKYVWLVEG 491
QY 494 HIDPNSKGLALASTLYGYN-----SNTIWRSMWMDNEVAFNNGSGSGDGDKPVPBPQD 549
DB 492 HYDNTGKN---LKTQVLOENVDPVTRNDYIFGWNENNVRYGGGSADG--DSAVNPKDP- 546
QY 550 EPGIEPIPEDSDSDPGSDSGSDSNDSDSGSDSTSDSGSDSASDSDSASDSDS 609
DB 547 TPG--PPVDPEPDP-----EPEPTPDP 568
QY 610 DSASDSDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 669
DB 569 EPGPDPPEPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDP 628
QY 670 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 729
DB 629 ESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 688
QY 730 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 789
DB 689 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 748
QY 790 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 849
DB 749 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 795
QY 850 SPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 908
DB 796 -SDSDSDSDSDR---VTPPNNEQKAPSNPKGEVNHNSKVSQHKTKDALPETGKSENTNAT 851
QY 909 LIWGLLASIGSLILFRKKENKOKK 933
```

Db 852 LFGAMMALLGSLILFRKKQDHKEK 876

RESULT 4
T30214

fibrinogen-binding protein - Staphylococcus epidermidis
C;Species: Staphylococcus epidermidis

C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C;Accession: T30214

R;Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66, 2666-2673, 1998

A;Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.

A;Reference number: Z20781; MUID:98261511; PMID:9596732

A;Accession: T30214

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1092 <N1>

A;Cross-references: UNIPROT:O70022; EMBL:Y17116; NID:el296734; PID:el296735; PIDN:CAA7663

Query Match 34.5%; Score 1628; DB 2; Length 1092;

Best Local Similarity 36.8%; Pred. No. 5e-65;

Matches 421; Conservative 141; Mismatches 297; Indels 284; Gaps 32;

QY 5 KKEKHAIRKKSIGVASVLVGTGILGGLSSKEADASENSV-----TQSDASN 52

DB 18 KSNKYAIRKFTVGTASIVIGATLLFG-LGHNEAKAEENSVQDKDSNTDDELSDSDQSS 76

QY 53 ESKSNSSSVSAAPKTDITNV-----SDTKTSNTN---NGETSVAQNP 93

DB 77 DEEKNNVINNQSGINTDDNNQIIKKEETNNYDGIKRSSEDTSTTNNVDENEATFLQKTP 136

QY 94 -----AQOETTQSSSTNATTEPTVGTGATTTTTTQANTPATQOS-SNTNAELVN--- 143

DB 137 QDNTHLTEEEVKSSSVSSNSSSIDTAQSPSHITINREESVQTSNDVEDSHVDFANSKI 196

QY 144 QTSNETTFPDNTV-----SSVNSPQNSNAENVSTTQDTSTTEATPSN---NESAP 191

DB 197 KESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN-LPINEVENKARP 255

QY 192 QSTDASNKDVVNOAVNTSAPRMRAFSLAAVAADAPAGTDTITNLTNVTYGI-DSGTTVY 250

DB 256 LSTTSAGPSIKRYTVN-----QLAAEQGSNVAHLIKVTDQ---SITEGYDDSEGVK 304

QY 251 PHOAGVVKLNYGFSVPSNAVKGTFTKITVPKELNLNGVTSTAKVPPIMAGD-OVLANGVI 309

DB 305 AHDENLIYDVTTEVDKVKSGDTMTVDIDKNVPSDLTDSFTPIPKIKNSGRIIATGY 364

QY 310 DS-DGNVIYFTFDYVNTKDDVKATLTMPAYIDPENVKKTG---NVTLATGIGSTTANKTV 365

DB 365 DNKNKQITVYFTDYVKYENIKAHLLKITSYDKSVNNNTKLDVEYKTALSS--VNKTI 422

QY 366 LVDPYKYGKFEYNLSIKGTTDQDKTNTYQTIYVNPNGDNVIAPLVTGLNKLKNTDSNAL 425

DB 423 TVEYQRPNERNTANLQSMFTNIDTKHTEVQTIYINPL--RYSAKETNWNISNGDEGST 480

QY 426 IDQOQNTSIKVKYVDNAADLSESYFV-NPENFEDVTNSVNITFPNPQYKVFENTPDDQIT 484

DB 481 IIDSDTIKVKYKGDQNLNPDNSRIYDYSEYEDVTDDYAQLGNNDVNINFG-----NID 536

QY 485 TPYIVVWNGHIDPNSKGD-----LALRSTLYGYANSIIWRSMWMDNEVAFNNGSGSD 537

DB 537 SPVILKISKYDEN-KDDYTTIQQVTVMQTTINEYGE--FRTASYDNTIAFTSSGQGG 593

QY 538 G-----IDKPVVP-----EQP-----DEPEIE 555

DB 594 GDLPEPKTYKIGDYVWEDVDKGIQNTNDNEKLSNVLVTLYPDGTSKSVRTDEGKYQ 653

QY 556 -----PIPEDSDSDPGSDSGSDS-----NSDSGSDG----- 582

DB 654 FDGLKNGLYTKITFTPEGTPTPLKHSNTINPALDSEGNVWVWINGODMTIDSGFYQTP 713

QY 583 ----- 582

C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A: title: Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*.
A: Reference number: A89758; MUID: 21311952; PMID: 11418146
A: Accession: D89824
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-1385 <KUR>
A: Cross-references: UNIPROT: Q99W47; GB: BA000018; PID: g13700454; PIDN: BAB41751.1; GSPDB: G13700454
A: Experimental source: strain N315
C: Genetics:
A: Gene: *sdhD*

Query Match 29.5%; Score 1395.5; DB 2; Length 1385;
Best Local Similarity 29.6%; Pred. No. 1.4e-54;
Matches 416; Conservative 149; Mismatches 329; Indels 513; Gaps 40;

QY 1 MNMKKEKHAIRKSGVASVLGTLIGFGL--LSSKEADASENSTQSDASNEKSNDS 58
DB 15 MVSNRLNKFIRKYVTGTASILVGTLLIFGLGQEAKEAESTKELNEATTSASDNQSSD 74
QY 59 SSSVSAAPKTDGT-NVSDTKTSNTNNGETSVQAQOFTTQSSSTNATTEETPTVGA 117
DB 75 KVDQMQLNQEDNTKDNQKEMWSQOGETTSNGKIEKESVOSTTGNKV-----EV 126
QY 118 TTTTTCQANTPATQSSNTNAELVNOTSNETTFTDNTNTVSSV--NSPONSTNAENVST 174
DB 127 STAKSDEQASPKSTNEDLNTKQIISQEGGLPDLLENKSVNVQPTNEENKKYDKATEST 186
QY 175 TQDTSTEAFTSNNEBAPQSTDA--KQVNVQAVNTSAPRRAFSLAAVAAD 224
DB 187 TLNVKSDAIKSAETLVDNNSNENNENNAIILPKSTAPKSLNT--RMRMAAIQPNSTD 243
QY 225 APAAGTDITNQLNVTVTGIDSGTIVVPHQAGYKVLNAGRSVPNSAVKGDFTKITVPKELN 284
DB 244 SKNVNDLITNTLTVVVDADNSKTIIPAQ--DYLSLSQITVDDKVKSGDYFTIKYSDTVQ 302
QY 285 LKGV-----STAKVPPIPMAGDQVLANGVIDSDGNVI--YTFTDYVNTKDDVKATLTMP 336
DB 303 VYGLNPEDIKNGIDKP---NNGETIATAKHDTANLLTYTFTDYVDRENSVKMGINYS 359
QY 337 AYIDPENV---KKTGNVTLAGTSTANKTVLVDYEKYGKFNLSIKG---TIDQIDK 389
DB 360 IYMDADTIPVDKK--DVPEFSTVIGNQITTTTADITYPAYKEADNNSIGSAFTVSHVGN 417
QY 390 TNNT--YRQFIYVNPBGDNVI-APVLGNLKPNTDSN-ALIDQNTSIIKVKYVDNADLS 445
DB 418 VEPGYNQVYVNPMDKLGAKLKVAEVHPKYPTNIGQINQNVNTNIKIYRPEGYTLN 477
QY 446 ESYFVNPENPEDVTN-----SVNITF----- 466
DB 478 KGVDTNLDVDTDEFKNKMTYGSQSNVLDGDI TSAYVMVNTKQYNTSESFTLVQ 537
QY 467 ----- 466
DB 538 MATLSSTGNKSVSTGNALGFTNNQSGAGQEVYKIGNYVWEDTNKNGVQELGKGVGNVT 597
QY 467 -----PN-PN-QYKVEF-NTPDDQIITTP----- 486
DB 598 VTFVDNNTTKVGEAVTKEDGSYLIPNLPGDYRVEFSNLPKGYEFTPSKQGNNEELDSN 657
QY 487 ---YIVVNGHDPNSKGDILALRSLTYGYSNIIWRSMQWDEVAFNNGSGSD--GIDK 541
DB 658 GLSSVITVG--KDNLSADLGIYKPKYNL-GDYVWEDTN-----KXGLQDQDEKIGS 707
QY 542 PVVPEQDFPEGEIPIPEDSDS-----DPGSDSGSDSN 574
DB 708 VTVTLKDENGNLKTVTTDADGKYKFTDLNGNKKVEFTTPEGYTPPTVTSGGSDIEKDSN 767
QY 575 -----SDGSDSGSDSTS----- 587
DB 768 GLTTGTGIVGADNMTLDSGFYKTPKYNLGNVYWEDTNKDGKQDSTEGKISGVTVTLKNE 827

[illegible]

R; Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A; Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mä
ok, C.; Schluerer, T.; Simoes, N.; Tierrez, A.; Vazquez-Roland, J.A.; Voss, H.; Wehländ,
A; Title: Comparative genomics of *Listeria* species.
A; Reference number: AB1077; MUID:21537279; PMID:11679669
A; Accession: AG1671
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-882 <GLA>
A; Cross-references: UNIPROT:Q92AK9; GB:AL592022; PIDN:CAC57143.1; PID:gl6414414; GSPDB:G
A; Experimental source: strain Clip11262
C; Genetics:
A; Gene: lin1913

```
Query Match      20.3%; Score 957.5; DB 2; Length 882;  
Best Local Similarity   32.1%; Pred. No. 2.e-35;  
Matches 178; Conservative 187; Mismatches 126; Indels 63; Gaps 9;
```

Qy		345 KKTGNVTILATIGTSTANKT--VLVDYKYGKFYNLS-IKGTIDIQDKNTNTRYQTIVVN	401
Dd	:	: : : : :	:
	:	: : : :	:
Dd	16 KRAQAKALIAALTITTTMMVAPVVNYDSFNHFKFALSIOADAATI DLLGNSSLNQZAN	75	
Qy	402 PSGDNVIAPVLTGNIKPNTPDSSNALIDQQNTSIHKYYKDVAADLSESYFYNPENFEDEVTNIS	461	
Dd	76 -----GKLVTLSGNQLVSA-----SAVSTYYPYFPFLPSELTSILSN	112	
Qy	462 VNITPPNNQYKVE-----ENTPDQIITTFPIVVVNG----HIDPSNKGDGLARLS	507	
Dd	113 PNIRANTIKDIKYAVLGIGNIGLFNOG-----TVNGSNNFFIDITSRNAIGARN	162	
Qy	508 TLYGIN-SNIWRMSWD-----NEVAFNMGSGDGIDKPVPVPEQPDEPGEI	554	
Dd	163 HLLGVGSVSFTLTIDLALGVTLPAANDCKLDFAARTGDLGVLDLLNSNAARGFI	222	
Qy	555 EPIPEDSDPGSDSGSDSNSGSDSGSDGSTSDSGSDASDSDSDSASDSDSASD	614	
Dd	223 TTVDGDAAD	282	
Qy	615 SDSASDSDMDNDSD	674	
Dd	283 AD	342	
Qy	675 SDSPSD	734	
Dd	343 AD	402	
Qy	735 SDSPSD	794	
Dd	403 AD	462	
Qy	795 SDSPSD	854	
Dd	463 AD	522	
Qy	855 SESDSNSDESNGSN	868	
Dd	523 ADADADADADADAD	536	

RESULT 12
F90073
hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F90073
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1235-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A;Reference number: A89758; MUID:213111952; PMID:11418146

A;Accession: F90073

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-2271 <KUR>

A;Cross-references: UNIPROT:Q99QY4; GB:BA000018; PID:gl3702612; PIDN:BAB43752.1; GSPDB:G

A;Experimental source: strain N315

C;Genetics:

A;Gene: SA2447

Query Match 16.9%; Score 799; DB 2; Length 2271;

Best Local Similarity 23.9%; Pred. No. 7.9e-28;

Matches 295; Conservative 170; Mismatches 406; Indels 116; Gaps 20;

Qy 12 RKKGIGVASVLVGLIGLLSSKEADASENSVTVQSDASNEKSNDSSSVSAAPKDDT 71
Db 1327 KSDSISTSLSGS-----TSESDSTSSSEKSDSTGMSISMSQSTGSTST-ST 1379
Qy 72 NVSD---TKTSSNTNNGETSAQNPAAQOETTPQSSSTNATTEETPTVGEATTTTNQANTP 128
Db 1380 SLSDSTSLSLASMMQSGVDNSASQASNSTSTSESDS---QSTYTSQSTSQ 1435
Qy 129 ATTQSSNTNAELVNQSTNETTFNDNTVSVSPQNSNAENVSTTQDTSTTEATPNNE 188
Db 1436 SESTSTSLSDSTSLGKSTSGSGSTSTASLSGSESDSQSTSTSEASTSL 1495
Qy 189 SAPOSTDASNKDVVNQAVNTSAPRMRAFLAAVAADAPAA---CTDITNLTNTVTGIDS 245
Db 1496 SDSTSTNSGASSTSLSNSASASESDSSSTSLSDSTASMQSESDSQSTSLSDSL 1555
Qy 246 GTTVYPHQAGYVKLYGFSVPNSAVKDTFKITVPKELNLNGVTSTAKVPPIMAGDOVLA 305
Db 1556 STSTSNRMSTIASLSTSVTSESGSTSESTSESDSTSLSDSQSTSR-SFSASGSASTS 1614
Qy 306 NGVIDSGNVIYTFDY-VNFKDDVKATLT-----MPAVIDPENVKKT 347
Db 1615 TSTDSTSTASTSTSMRTSTSDSQMSLSTSTSTSMSDSTSLSDSVSDSTSDSTASTS 1674
Qy 348 GNVTLATIGSTTANKTVLDYKYGKFNLSIKGTIDQIDKNTYRQTIYVNPSSGDNV 407
Db 1675 GSMVSLSLSDSTSTSTASVMGASISDQSMSESVNDSSESVSESDSKMSGSGST 1734
Qy 408 IAPVLITGNLKENT-----DSNALDQNTIKYVKVDNAADLSESYFNPENFEDV 458
Db 1735 VSD--SGSLSVSTSLRKSSESVESGLSGQMSDVSSTSDS--SLSVSTSLRSSESV 1789
Qy 459 TNSVNIPTPNQYKVFENFPDDQITTPYIVVNGHIDPNKGDALRSTLYGYNSNIW 518
Db 1790 SESDSL-----DSKST-----SGSTSTSGSLSTSTSLG--SESVS 1826
Qy 519 RSMGMDNEVAFNNGSGSD-----GIDKPVVPEQDEPEGEIPEIPEDSDSPGSDSG 570
Db 1827 ESTSLSDSISMSDSTSTSDSLSGSTSLSGSTSLSTSDSLSDSKSL-SSSQSMGSGEST 1885
Qy 571 SDSNSD 630
Db 1886 STSVSDSQSSSTSNQPDMSGISASESDSMGTSDDSNISGNSSTSLSTSDSMGSGSV 1945
Qy 631 SD 690
Db 1946 STSTSLSDSISGTSVSDSSSTSTSTSLSDSMQSQSTSTSTASGSLSTSTSTSMGSGAST 2005
Qy 691 SD 750
Db 2006 SSSQSTSVSTSLSDSTSLSDSTSLISGQSTVSESESTSDSTSLSDSESLSTSDSDSTST 2065
Qy 751 SD 810
Db 2066 STSDSTSGSTSTSLSESLSTSGSGSTSVSDSTSMSESDSTSVMSQDKSDSTSLSDSESV 2125
Qy 811 SDSESD 859
Db 2126 STSTSTSLSTSDSTSTSESLSTSMGSGSQISDS*STSMGSGSTSTSESNMHPDSMSMH 2185

Qy 860 -----NSDSESGSNVVPNPNKNGTNASNKNEAKDSKEPLPDTGSEDEA 905
Db 2186 THSTSTSRLSSEATTSTSESQSTLSATSEVTKNGTGA-----QSEKRLPDTGDSIKQ 2238
Qy 906 NTSILWGLLASIGSLLFRKKENKOK 932
Db 2239 N-GLGGVMTLLVGLGLMKRKK-KDE 2263

RESULT 13

G90053

hypothetical protein fnbB [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C;Accession: G90053

R;Kuroda, M.; Ohra, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A;Reference number: A89758; MUID:213111952; PMID:11418146

A;Accession: G90053

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-961 <KUR>

A;Cross-references: UNIPROT:Q99RD3; GB:BA000018; PID:gl3702452; PIDN:BAB43593.1; GSPDB:G

A;Experimental source: strain N315

C;Genetics:

A;Gene: fnbB

Query Match 14.6%; Score 688.5; DB 2; Length 961;

Best Local Similarity 22.6%; Pred. No. 2.1e-23;

Matches 236; Conservative 176; Mismatches 428; Indels 205; Gaps 29;

Qy 3 MKKKEKHAIRKKSIGVASVLVGLTIGLLSSKEADASE---NSVTQSDASNEKSNDS 59
Db 1 MKNLRYGIRKHKLGAASVFLGTMIVVGMQGEKAAASEQNNTTVEESGSSATESKA--- 57
Qy 60 SSVSAAPKTDNTVSDTKTSSNTNNGETSAQNPAAQOETTPQSSSTNATTEETPTVGEATT 119
Db 58 -----SETQTTNNNTTIDETQSYSATSTEQPSKSTQVTTTEAPTQVQPK 103
Qy 120 TTNQANTPATQSSNTNAELVNQSTNETTFNDNTVSVSPQNSSTNAENVSTTQDTS 179
Db 104 VETEMKS-----QEDLPSEKVDKETTQ 126
Qy 180 TE---ATPSNNEAPOSTDASNKDVVNQAVNTSAPRMR-----AFSLAAVAADAPAAGT 230
Db 127 TVQDIAQPSN-----VSEIKPRMKRGADVTAVGEKEVAEEAKATGT 167
Qy 231 DITNLTNTVTGID----SGTTVYPHQAGYVKLYGFSVPNSAVKDTFKITVPKELNLN 286
Db 168 DVTNKKEVTESSLEGHKNKDSNIVNPHNAQRVTLTKYKKGEGIKAGDYFTDLSDNVETH 227
Qy 287 GVTSTAKVPPIMAG--DQVLANGVIDSDGNVIYTFDYVNTKDVKATLTWPAYIDPENV 344
Db 228 GISTLRKVPKELKSTEDKMWANGQVNERIRYTFDYINNKKDLTAELNLNLFIDTTV 287
Qy 345 KKTGNVTLATIGGSTTANKTVLDY-----EKYGFNLSIKGTIDQIDKNTNTRYOTIY 399
Db 288 TKQSGQKQVETLQGNKYVSKFEDIKYLDGVKDRMG---VTVNGRIDTLNKEEGKFSHFAY 343
Qy 400 VNPSGDNIAPVLTCNLKPNKPTDSNALIDQOQNTSKYKVDNADLSSEYFV---NPNFE 456
Db 344 VKPNQSLTSTVTGQVTSQYKQSA-----NNPTVKYKHKHGSDELAESVAKLDDDTSKFE 399
Qy 457 DVTVSNVITPFPNQYKVEENTPDQITTPYIVVNGHIDPNKGDALRSTLYGNSNI 516
Db 400 DVTKEVNLSTNSGGYTLNLGDLN--SKDYVIKYEGEYQONAK-DLNFRTLHLSGYHKY 456
Qy 517 IWR-----SMSWDNEVAFNNGSGSDGIDKPVPW-----EQDPEGEI 554

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2004, 04:52:14 ; Search time 1562 Seconds

(without alignments)
11755.746 Million cell updates/sec

Title: US-09-679-643-1

Perfect score: 3498

Sequence: 1 ggtaccataaattacacatc.....gaagagtataaagaaagctt 3498

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 23Sep04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3498	100.0	3498	5	Aaf58593 Staphyloc
2	3487	99.7	3499	3	Aaz55832 Staphyloc
3	3279.6	93.8	4709	2	Aav74898 Staphyloc
4	2797.4	80.0	2799	8	Abt14896 Pathogen
5	2746.6	78.5	2781	8	AcF73781 Staphyloc
6	2331.6	66.7	2808	8	ACA20272 Prokaryot
7	2096.8	59.9	3066	4	Aas54809 Staphyloc
8	2093.8	59.9	3063	4	AAS51834 Staphyloc
9	1776.4	50.8	2792	6	Abk13616 S. aureus
10	1776.4	50.8	2792	8	ADA89855 Staphyloc
11	1560	44.6	1560	6	Aad46861 Staphyloc
12	990	28.3	990	6	Aad46862 Staphyloc
13	619.2	17.7	5406	3	Aaz93533 Cell wall
14	619.2	17.7	5406	3	Aaz51201 Staph. ep
15	589.6	16.9	4627	4	Aah54581 S. epider
16	589.2	16.8	4899	8	ACA46383 Prokaryot
17	584.8	16.7	1485	2	Aah52321 S. epider
18	568.6	16.3	2369	8	Aax77591 S. aureus
19	568	16.2	2739	2	Abt14866 Pathogen
20	558.2	16.0	1059	6	AbS52617 Staphyloc
21	532.4	15.2	1539	8	ACA35100 Prokaryot

22	525.4	15.0	2151	8	ABT14861	Pathogen
23	510	14.6	2631	8	ACF74262	Staphyloc
c	510	14.6	3007	8	ADA89797	Staphyloc
25	506.4	14.5	2010	8	ACA46390	Prokaryot
c	502.6	14.4	2582	8	ADA89798	Staphyloc
c	502.6	14.4	2582	8	ADA89868	Staphyloc
28	502	14.4	2634	8	ACA46374	Prokaryot
29	497.8	14.2	11050	2	AAV74407	Staphyloc
30	485.6	13.9	4911	8	ABT15076	Pathogen
31	462	13.2	3600	8	ABT15076	Staphyloc
32	448.4	12.8	2985	8	ACF73997	Staphyloc
c	446	12.8	2428	8	ADA89818	Staphyloc
34	439.4	12.6	4158	8	ACA20270	Prokaryot
c	436.8	12.5	756	4	AAH52319	S. epider
36	436.6	12.5	4143	8	ABT14918	Pathogen
37	415	11.9	5764	8	ACA47359	Prokaryot
c	383.6	11.0	4418	2	AAV74765	Staphyloc
39	375.6	10.7	2796	4	AAS51941	Staphyloc
40	375.6	10.7	2799	4	AAS54703	Staphyloc
c	365	10.4	396	2	AAV77913	Staphyloc
42	361.4	10.3	2841	8	ABT14889	Pathogen
43	360.6	10.3	3498	8	ABT14929	Pathogen
c	360.6	10.3	4005	2	AAV74942	Staphyloc
45	360.6	10.3	4047	8	ACF73996	Staphyloc

ALIGNMENTS

RESULT 1
AAF58593
ID AAF58593 standard; DNA; 3498 BP.
XX
AC AAF58593;

DT 23-APR-2001 (first entry)

DE Staphylococcus aureus clfa gene.

KW Staphylococcus aureus; clfa; antibiotic; vaccine;
XX fibrinogen binding protein; bacterial infection; mastitis; ds.

OS Staphylococcus aureus.

PN US6177084-B1.

PD 23-JAN-2001.

PF 19-OCT-1999; 99US-00421868.

PR 22-AUG-1994; 94US-00293728.

XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

PI Foster TJ, Mcdevitt DL;

DR WPI: 2001-181608/18.

XX P-PSDB; AAB69508.

PT Novel Staphylococcus aureus fibrinogen binding protein useful as vaccine
PT to prevent infection, promote wound healing, block adherence to
PT indwelling medical devices and for diagnosing staphylococcus aureus
PT infection.

XX Claim 1; Fig 2; 30pp; English.

CC The present sequence encodes a novel Staphylococcus aureus fibrinogen
CC binding protein. It is useful as a vaccine to protect against human and
CC animal infections caused by S. aureus, such as against mastitis, to block
CC S. aureus from colonising and infecting a wound, to block adherence of
CC S. aureus to indwelling medical devices such as catheters, replacement
CC heart valves and cardiac assist devices. The protein can be used to
CC diagnose bacterial infections. Polyclonal and monoclonal antibodies

Oppi, car

CC raised against the fibrinogen binding protein are useful for passively
CC immunising against infections caused by S.aureus, to prevent infection of
CC a wound and to diagnose bacterial infections

XX	Sequence	3498 BP; 1260 A; 587 C; 704 G; 947 T; 0 U; 0 Other;	
Qy	Query Match	100.0%; Score 3498; DB 5; Length 3498;	
Db	Best Local Similarity	100.0%; Pred. No. 0;	
Qy	Matches 3498; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	GGTACCAATAATACATCTGCTTTTGAAGAAATATGATTTCAAGTAGGATACATTA	60
Db	1	GGTACCAATAATACATCTGCTTTTGAAGAAATATGATTTCAAGTAGGATACATTA	60
Qy	61	GGTAGAGTTTCATATTAATAATAAAAAATGTTGCAATCAATCGTAGCTTGTCTTGTGA	120
Db	61	GGTAGAGTTTCATATTAATAATAAAAAATGTTGCAATCAATCGTAGCTTGTCTTGTGA	120
Qy	121	ATTCTTAAAAATAGCAATTAATAATGTTTGTAGTAAAGTATTAATTTGGGATATAAAA	180
Db	121	ATTCTTAAAAATAGCAATTAATAATGTTTGTAGTAAAGTATTAATTTGGGATATAAAA	180
Qy	181	TATCGATACAAATTAATGCTATTAATGCAATTTTAGTGATATAATTCATTAACAGAGTT	240
Db	181	TATCGATACAAATTAATGCTATTAATGCAATTTTAGTGATATAATTCATTAACAGAGTT	240
Qy	241	AAATATATCTTAAAGGCTATATAGTTAATATAAATGACTTTAAAGAGAGGGAATAAA	300
Db	241	AAATATATCTTAAAGGCTATATAGTTAATATAAATGACTTTAAAGAGAGGGAATAAA	300
Qy	301	ATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTTGGCGTGGCTTCA	360
Db	301	ATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTTGGCGTGGCTTCA	360
Qy	361	GTGCTTTAGTAGTACGTTAAATCGGTTTTGGACTACTACAGAGTAAAGAGCAGATGCAAGT	420
Db	361	GTGCTTTAGTAGTACGTTAAATCGGTTTTGGACTACTACAGAGTAAAGAGCAGATGCAAGT	420
Qy	421	GAAATATGTTACGCAATCTGATAGCGCAAGTAACGAAAGCAAGCAAAAGTAAATGATTCAGT	480
Db	421	GAAATATGTTACGCAATCTGATAGCGCAAGTAACGAAAGCAAGCAAAAGTAAATGATTCAGT	480
Qy	481	AGCGTTAGTGTGACCTCAAAACAGACACAAACGTTAGTGATCTATAAAACATCGTCA	540
Db	481	AGCGTTAGTGTGACCTCAAAACAGACACAAACGTTAGTGATCTATAAAACATCGTCA	540
Qy	541	AAACACTTAATGCGGAAACGAGTGTGCGCAAAATCCAGCACACAGGAAACGACACAA	600
Db	541	AAACACTTAATGCGGAAACGAGTGTGCGCAAAATCCAGCACACAGGAAACGACACAA	600
Qy	601	TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGTTGAAGCTACTACTACG	660
Db	601	TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGTTGAAGCTACTACTACG	660
Qy	661	ACAAACGAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA	720
Db	661	ACAAACGAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA	720
Qy	721	TTAGTGAATCAACAAAGTAAAGCACTTTTAATGATTAATTAATCAAGTATCATCTGTA	780
Db	721	TTAGTGAATCAACAAAGTAAAGCACTTTTAATGATTAATTAATCAAGTATCATCTGTA	780
Qy	781	AAITTCACCTCAAAATCTCAAAATGCGAAAAATGTTTCAACACGCAAGTACTTCAACT	840
Db	781	AAITTCACCTCAAAATCTCAAAATGCGAAAAATGTTTCAACACGCAAGTACTTCAACT	840
Qy	841	GAAGCAACACCTTCAACAAATGATAGCTCCACAGAGTACAGATGCAAGTAAATAAGAT	900
Db	841	GAAGCAACACCTTCAACAAATGATAGCTCCACAGAGTACAGATGCAAGTAAATAAGAT	900
Qy	901	GTAGTTAATCAACGGGTTAATACAAAGTGGCGCTTAGATGAGAGATTTAGTTTACGGCA	960
Db	901	GTAGTTAATCAACGGGTTAATACAAAGTGGCGCTTAGATGAGAGATTTAGTTTACGGCA	960

Qy	961	GTAGTGCAGATGCACCGCAGCTGGCACAGATATTACGAATCAGTTGACGAATGTGACA	1020
Db	961	GTAGTGCAGATGCACCGCAGCTGGCACAGATATTACGAATCAGTTGACGAATGTGACA	1020
Qy	1021	GTTGGTATTGACTCTGGTAGCACTGTGTATCCGCCACCAAGCAGGTTATGTCAAACTGAAT	1080
Db	1021	GTTGGTATTGACTCTGGTAGCACTGTGTATCCGCCACCAAGCAGGTTATGTCAAACTGAAT	1080
Qy	1081	TATGGTTTTTTCAGTGGCCTAATTTCTGCTGTTTAAAGGTGACACATTCAAAATTAACCTGACCT	1140
Db	1081	TATGGTTTTTTCAGTGGCCTAATTTCTGCTGTTTAAAGGTGACACATTCAAAATTAACCTGACCT	1140
Qy	1141	AAAGAAATTAACCTTAAATGGTAACTTCAACTGCTAAAGTGCACCACTTATGGCTGGA	1200
Db	1141	AAAGAAATTAACCTTAAATGGTAACTTCAACTGCTAAAGTGCACCACTTATGGCTGGA	1200
Qy	1201	GATCAAGTATTGGCAATGGTAAATCGATAGTAGTAAATGTTTATTAATACATTTACA	1260
Db	1201	GATCAAGTATTGGCAATGGTAAATCGATAGTAGTAAATGTTTATTAATACATTTACA	1260
Qy	1261	GACTATGTAATTAATAAGATGATGTAAGAGCACTTTGACCATGCCCGCTTATATTGAC	1320
Db	1261	GACTATGTAATTAATAAGATGATGTAAGAGCACTTTGACCATGCCCGCTTATATTGAC	1320
Qy	1321	CCTGAAAAATCTTAAAAAGACAGGTAATGTCACATTTGGCTACTGGCATAGGTAGTACAACA	1380
Db	1321	CCTGAAAAATCTTAAAAAGACAGGTAATGTCACATTTGGCTACTGGCATAGGTAGTACAACA	1380
Qy	1381	GCAAAACAAAACAGTATTAGTAGTAAATGTAAGTAAATGTTTAACTTATCTATT	1440
Db	1381	GCAAAACAAAACAGTATTAGTAGTAAATGTAAGTAAATGTTTAACTTATCTATT	1440
Qy	1441	AAAGGTACAAATTCACCAATTCGATAAACAATAATACGTTATCGTCAGACAAATTTATGTC	1500
Db	1441	AAAGGTACAAATTCACCAATTCGATAAACAATAATACGTTATCGTCAGACAAATTTATGTC	1500
Qy	1501	AAATCCAAAGTGCAGATACGTTATTGGCGCGGTTTAAACAGGTAATTTAAACCAAAATACG	1560
Db	1501	AAATCCAAAGTGCAGATACGTTATTGGCGCGGTTTAAACAGGTAATTTAAACCAAAATACG	1560
Qy	1561	GATAGTAATGCAATTAATAGATCAGCAAAATACAAATGTAATAAGTAGATAAT	1620
Db	1561	GATAGTAATGCAATTAATAGATCAGCAAAATACAAATGTAATAAGTAGATAAT	1620
Qy	1621	GCAGCTGATTTATCTGAAGTTTACTTTGCAATCCAGAAACCTTTGAGGATGTCACCTAAT	1680
Db	1621	GCAGCTGATTTATCTGAAGTTTACTTTGCAATCCAGAAACCTTTGAGGATGTCACCTAAT	1680
Qy	1681	AGTGTCAATATTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATACGCTGAT	1740
Db	1681	AGTGTCAATATTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATACGCTGAT	1740
Qy	1741	GATCAAAATTAACAACCGTATATAGTAGTTGTTTAAATGGTCAATATTGATCCGAATAGCAAA	1800
Db	1741	GATCAAAATTAACAACCGTATATAGTAGTTGTTTAAATGGTCAATATTGATCCGAATAGCAAA	1800
Qy	1801	GGTGATTTAGCTTTTACGTTCAACTTTATATGGGTATAAATCGAATATAATTTGGCGCTCT	1860
Db	1801	GGTGATTTAGCTTTTACGTTCAACTTTATATGGGTATAAATCGAATATAATTTGGCGCTCT	1860
Qy	1861	ATGCTCATGGGACACCAAGTAGCATTTAATAACGGATCAAGTTCTGGTGACCGGTATCGAT	1920
Db	1861	ATGCTCATGGGACACCAAGTAGCATTTAATAACGGATCAAGTTCTGGTGACCGGTATCGAT	1920
Qy	1921	AAACCAAGTTGTTCTGAAACCAACCTGATGAGCTGGTGAATTTGAAATTCAGAGGAT	1980
Db	1921	AAACCAAGTTGTTCTGAAACCAACCTGATGAGCTGGTGAATTTGAAATTCAGAGGAT	1980
Qy	1981	TCAGATTTCTGACCCAGGTTTCAGATTTCTGGCAGAGGATTTCTAAATTCAGATACGGGTTACAGAT	2040
Db	1981	TCAGATTTCTGACCCAGGTTTCAGATTTCTGGCAGGCAATTTCTAAATTCAGATACGGGTTACAGAT	2040

2041 TCGGTAGTGTCTTACATCAGATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGCA 2100
Db
2041 TCGGTAGTGTCTTACATCAGATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGCA 2100
Qy
2101 AGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGAT 2160
Db
2101 AGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGAT 2160
Qy
2161 TCGGACTCAGCAATGACTCCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2220
Db
2161 TCGGACTCAGCAATGACTCCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2220
Qy
2221 TCGGACAGTGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2280
Db
2221 TCGGACAGTGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2280
Qy
2281 AGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2340
Db
2281 AGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2340
Qy
2341 TCTGACTCCGACAGTGATTCGACTCCGACTCCGACTCCGACTCCGACTCCGACTCCGACT 2400
Db
2341 TCTGACTCCGACAGTGATTCGACTCCGACTCCGACTCCGACTCCGACTCCGACTCCGACT 2400
Qy
2401 TCAGATAGCGATTCGACTCCGACTCCGACTCCGACTCCGACTCCGACTCCGACTCCGACT 2460
Db
2401 TCAGATAGCGATTCGACTCCGACTCCGACTCCGACTCCGACTCCGACTCCGACTCCGACT 2460
Qy
2461 AGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2520
Db
2461 AGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2520
Qy
2521 TCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2580
Db
2521 TCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2580
Qy
2581 TCAGACAGTGATTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2640
Db
2581 TCAGACAGTGATTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2640
Qy
2641 AGTGACTCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2700
Db
2641 AGTGACTCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2700
Qy
2701 TCGGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2760
Db
2701 TCGGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2760
Qy
2761 TCAGACAGCGACTCAGACAGTGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2820
Db
2761 TCAGACAGCGACTCAGACAGTGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2820
Qy
2821 TCAGACTCAGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2880
Db
2821 TCAGACTCAGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2880
Qy
2881 AGCGATTCGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2940
Db
2881 AGCGATTCGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2940
Qy
2941 ACTAATGCTCTTAATAAATAGGCTAAAGATAGTAAAGAACCATTTACAGATACAGT 3000
Db
2941 ACTAATGCTCTTAATAAATAGGCTAAAGATAGTAAAGAACCATTTACAGATACAGT 3000
Qy
3001 TCTGAAGATGAAGCAAAATACGTCATTAATTTGGGATTTATAGCATCAATAGGTTTCATTA 3060
Db
3001 TCTGAAGATGAAGCAAAATACGTCATTAATTTGGGATTTATAGCATCAATAGGTTTCATTA 3060
Qy
3061 CTACATTTTCAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3120
Db
3061 CTACATTTTCAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3120
Qy
3121 TAATCATATGATTCATGAAGAACCCACTTAAAGAGTGCTCTTTTACTTGGATTTTCCA 3180

Db 3121 TAATCATATGATTCATGAAGAACCCACTTAAAGAGTGCTCTTTTACTTGGATTTTCCA 3180
Qy 3181 AATATATGTTTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3240
Db 3181 AATATATGTTTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3240
Qy 3241 TAGATGTTTATATAATTTGGCTTGGCGAAAAAATAGGCTGTAAGGTAGGTGTTTAAATTAGG 3300
Db 3241 TAGATGTTTATATAATTTGGCTTGGCGAAAAAATAGGCTGTAAGGTAGGTGTTTAAATTAGG 3300
Qy 3301 GAAATTAAGGAGAAAAATACAGTTGAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3360
Db 3301 GAAATTAAGGAGAAAAATACAGTTGAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3360
Qy 3361 ATGTGTATCAAAATTTGGGAAAGTAATCGTGGAGTGCAGTGGTTCTGGGGAGAGAA 3420
Db 3361 ATGTGTATCAAAATTTGGGAAAGTAATCGTGGAGTGCAGTGGTTCTGGGGAGAGAA 3420
Qy 3421 TCCATATGATCTAGTCTGTTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 3480
Db 3421 TCCATATGATCTAGTCTGTTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 3480
Qy 3481 AGAGTATTAAGAAAGCTT 3498
Db 3481 AGAGTATTAAGAAAGCTT 3498

RESULT 2
AAZ55832
ID AAZ55832 standard; DNA; 3499 BP.

XX AC AAZ55832;
XX DT 27-MAR-2000 (first entry)
XX DE Staphylococcus aureus clfa gene.
XX KW Clfa; fibrinogen binding protein; bacterial colonisation;
XX KW indwelling medical device; staphylococcal infection; ds.
XX OS Staphylococcus aureus.

XX Key Location/Qualifiers
FT -35_signal 239..244 /*tag= a
FT -10_signal 261..266 /*tag= b
FT RBS 289..294 /*tag= c
FT CDS 302..3103 /*tag= d
FT /product= "Staphylococcus Clfa protein"

XX US6008341-A.

XX PD 28-DEC-1999.

XX PF 22-AUG-1994; 94US-00293728.

XX PR 22-AUG-1994; 94US-00293728.

XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

XX Foster TJ, Mcdevitt DL;

XX WPI; 2000-096389/08.

XX P-PSDB; AAY58435.

XX Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus

XX aureus, useful for treatment or prevention of infections.

XX Claim 8; Fig 2A-1-4; 35pp; English.

afp

XX This sequence represents the Staphylococcus aureus clfa gene which
CC encodes a fibrinogen-binding protein, Clfa. Clfa is an important receptor
CC involved in S. aureus colonisation of indwelling medical devices (e.g.,
CC catheters, artificial heart valves). Shortly after implantation, the
CC surfaces of medical devices become coated with host plasma and matrix
CC proteins such as fibrinogen and fibronectin, and there is considerable
CC evidence to suggest that bacterial adherence to fibrinogen/fibrin is
CC important in the initiation of device-related infection. The fibrinogen-
CC binding region of Clfa is thought to reside between residues 332 and 550
CC in a region designated A. The protein also contains a repeated region
CC (region R) comprising 154 repeats of the dipeptide Ser-Asp, and the C-
CC terminus contains features present in surface proteins of other Gram
CC positive bacteria that are responsible for anchoring the protein to the
CC cell wall and cell membrane. Clfa, or its fragments, may be used to block
CC S. aureus colonisation of wounds, to prevent adherence of S. aureus to
CC indwelling medical devices, as vaccines to protect against S. aureus
CC infection (e.g., mastitis in ruminants), to raise specific antibodies,
CC and for diagnosis (by agglutination or immunoassay). The specific
CC antibodies are used for passive immunisation, to block infection of
CC wounds or adhesion of S. aureus and for diagnosis. Nucleotides encoding
CC Clfa and its fragments may be used as diagnostic probes
XX

SQ Sequence 3499 BP; 1260 A; 587 C; 704 G; 948 T; 0 U; 0 Other;
Query Match 99.7%; Score 3487; DB 3; Length 3499;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3498; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGTACCAATAATACACATCTGCTTTTGAAGAAATATGATTTCAAGTAGGATACATTA 60
DB 1 GGTACCAATAATACACATCTGCTTTTGAAGAAATATGATTTCAAGTAGGATACATTA 60

QY 61 GGTAGCTTCATATTATATAAATAATGTTTGAATCAATCTAGCTGTTGCTTTGTA 120
DB 61 GGTAGCTTCATATTATATAAATAATGTTTGAATCAATCTAGCTGTTGCTTTGTA 120

QY 121 ATTCTTAAATAGCAATAAATAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA 180
DB 121 ATTCTTAAATAGCAATAAATAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA 180

QY 181 TATCGATACAAATAATGCTATAATGCAATTTTGTAGTAAAGTATTATTGTGGATAATAA 240
DB 181 TATCGATACAAATAATGCTATAATGCAATTTTGTAGTAAAGTATTATTGTGGATAATAA 240

QY 241 AAATATATC-TTAAAGGATATAGTTAATATAAATGACTTTTAAAGACAGGATAA 299
DB 241 AAATATATCTTTAAAGGATATAGTTAATATAAATGACTTTTAAAGACAGGATAA 300

QY 300 AATCAATATCAAGAAAAAAGAAAAACACGCAATTCGGAAAAATPCGATTCGGCTGCTC 359
DB 301 AATGAATATCAAGAAAAAAGAAAAACACGCAATTCGGAAAAATPCGATTCGGCTGCTC 360

QY 360 AGTGCCTGTAGTACGTTAATCGTTTGTGACTACTCAGCAGTAAAGACAGATGCAAG 419
DB 361 AGTGCCTGTAGTACGTTAATCGTTTGTGACTACTCAGCAGTAAAGACAGATGCAAG 420

QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAGCAAPAGTATGTTCAAG 479
DB 421 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAGCAAPAGTATGTTCAAG 480

QY 480 TAGCGTTAGTGTGCACTTAAACACAGACGACAAACGTCGAGTGATCTAAACATCGTC 539
DB 481 TAGCGTTAGTGTGCACTTAAACACAGACGACAAACGTCGAGTGATCTAAACATCGTC 540

QY 540 AATCACTAATAATGGCGAAACAGTGTGGCGAAATCCAGCACAAACAGGAACGACACA 599
DB 541 AATCACTAATAATGGCGAAACAGTGTGGCGAAATCCAGCACAAACAGGAACGACACA 600

QY 600 ATCATCATCAACAATGCACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 659
DB 601 ATCATCATCAACAATGCACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 660

QY 660 GACAAACGAATCAAGCTAATAACACGGCAACAACTCAATCAAGCAATACAAAATGGGAGGA 719
DB 661 GACAAACGAATCAAGCTAATAACACGGCAACAACTCAATCAAGCAATACAAAATGGGAGGA 720

QY 720 ATTAGTGAATCAAAACAAAGTAAATGAACGACTTTTAAATGATACATAATACAGTATCATCTGT 779
DB 721 ATTAGTGAATCAAAACAAAGTAAATGAACGACTTTTAAATGATACATAATACAGTATCATCTGT 780

QY 780 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACACGCAAGTATCTCAAC 839
DB 781 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACACGCAAGTATCTCAAC 840

QY 840 TGAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAAATAAGA 899
DB 841 TGAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAAATAAGA 900

QY 900 TGTAGTTAATCAAGCGGTTTAAATCAAGTGCCTTAGAATGAGAGCATTTAGTTAGCGGC 959
DB 901 TGTAGTTAATCAAGCGGTTTAAATCAAGTGCCTTAGAATGAGAGCATTTAGTTAGCGGC 960

QY 960 AGTAGCTGCAGATCCACCGGACGTCGACAGATATTACGAATCAGTTGACGATGTGAC 1019
DB 961 AGTAGCTGCAGATCCACCGGACGTCGACAGATATTACGAATCAGTTGACGATGTGAC 1020

QY 1020 AGTTGGTATTGACTCTCGTACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAACTGAA 1079
DB 1021 AGTTGGTATTGACTCTCGTACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAACTGAA 1080

QY 1080 TTATGGTTTTTCAGTGCCTAAATCTCTGCTGTAAAGGTGACACATTTCAAAATAACTGTACC 1139
DB 1081 TTATGGTTTTTCAGTGCCTAAATCTCTGCTGTAAAGGTGACACATTTCAAAATAACTGTACC 1140

QY 1140 TAAAGAAATTAACCTTAAATCGTGTAACTTCAACTGCTTAAAGTGCACCAATATATGGCTGG 1199
DB 1141 TAAAGAAATTAACCTTAAATCGTGTAACTTCAACTGCTTAAAGTGCACCAATATATGGCTGG 1200

QY 1200 AGATCAAGTATTGCGAAATCGTGTAACTCGATAGTGTGTAATGTTTATTTATACATTTAC 1259
DB 1201 AGATCAAGTATTGCGAAATCGTGTAACTCGATAGTGTGTAATGTTTATTTATACATTTAC 1260

QY 1260 AGACTATGTAATACTAAAGATGATGTAAGCAACTTTGACCAATGCGCGCTTATATGA 1319
DB 1261 AGACTATGTAATACTAAAGATGATGTAAGCAACTTTGACCAATGCGCGCTTATATGA 1320

QY 1320 CCTGGAATGTTTAAAGACAGTAAATGTCATTCGCTGCTTAAAGTGCACCAATATATGGCTGG 1379
DB 1321 CCTGGAATGTTTAAAGACAGTAAATGTCATTCGCTGCTTAAAGTGCACCAATATATGGCTGG 1380

QY 1380 AGCAAAACAAACAGTATTAGTATGTAATGAAAAATATGTAAGTTTATTAACCTTATCTAT 1439
DB 1381 AGCAAAACAAACAGTATTAGTATGTAATGAAAAATATGTAAGTTTATTAACCTTATCTAT 1440

QY 1440 TAAAGGTACAAATTCGCAAAATCGATAAACAATAATACGATATCGTCAGCAATTTATGT 1499
DB 1441 TAAAGGTACAAATTCGCAAAATCGATAAACAATAATACGATATCGTCAGCAATTTATGT 1500

QY 1500 CAATCAAGTGGAGATAACGTTATTCGCGCGTTTAAACAGGTAATTTAAACCAATATAC 1559
DB 1501 CAATCAAGTGGAGATAACGTTATTCGCGCGTTTAAACAGGTAATTTAAACCAATATAC 1560

QY 1560 GGATAGTAATGCAATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAA 1619
DB 1561 GGATAGTAATGCAATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAA 1620

QY 1620 TGCAGCTGATTTTCTGAAAGTACTTTGTGATCCAGAAACCTTTGAGGATGCTCACTAA 1679
DB 1621 TGCAGCTGATTTTCTGAAAGTACTTTGTGATCCAGAAACCTTTGAGGATGCTCACTAA 1680

QY 1680 TAGTGTGAATATTACATTCCTCAATCCAAATCAATTAAGTAGAGTTTAAATACGCGCTGA 1739
DB 1681 TAGTGTGAATATTACATTCCTCAATCCAAATCAATTAAGTAGAGTTTAAATACGCGCTGA 1740

QY 1740 TGATCAAAATTAACAACCGGATATATAGTGTGTTTAAATGTCATTTGATCCGAATAGCAA 1799

```
1741 TGATCAAAATTACACACCGTATATAGTAGTGTATATGCTCATATTTGATCCGAATAGCAA 1800
1800 AGGTGATTTAGCTTTACGTTTAACTTTATATGCGGTATATACTCGAATATATTTGGCGCTC 1859
1801 AGGTGATTTAGCTTTACGTTTAACTTTATATGCGGTATATACTCGAATATATTTGGCGCTC 1860
1860 TATGTCATCGGCAACGAAGTAGCATTTAATACCGATCAGGTTCTGGTGACGGTATCGA 1919
1861 TATGTCATCGGCAACGAAGTAGCATTTAATACCGATCAGGTTCTGGTGACGGTATCGA 1920
1920 TAAACCCAGTGTTCCTGGAACAACTGATGAGCTGCTGTAATTTGAACTTCCAGAGGA 1979
1921 TAAACCCAGTGTTCCTGGAACAACTGATGAGCTGCTGTAATTTGAACTTCCAGAGGA 1980
1980 TTCCAGATTCGACCCAGGTTTCAGATTCGCGAGCGATTTCTAATTCAGATAGCGGTTTCAGA 2039
1981 TTCCAGATTCGACCCAGGTTTCAGATTCGCGAGCGATTTCTAATTCAGATAGCGGTTTCAGA 2040
2040 TTCCGGTAGTGATTTACATCAGATAGTGTTTCAGATTCAGCGAGTGATTCAGATTCAGC 2099
2041 TTCCGGTAGTGATTTACATCAGATAGTGTTTCAGATTCAGCGAGTGATTCAGATTCAGC 2100
2100 AAGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGCGA 2159
2101 AAGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGCGA 2160
2160 TTCCGAGCTCAGACATGACTCCGATTCAGATAGCGGATTCGACTCAGACAGTGACTCAGA 2219
2161 TTCCGAGCTCAGACATGACTCCGATTCAGATAGCGGATTCGACTCAGACAGTGACTCAGA 2220
2220 TTCCGAGCTCAGACTCAGATTCAGATAGCGGATTCGACTCAGACAGTGACTCAGATTCAGA 2279
2221 TTCCGAGCTCAGACTCAGATTCAGATAGCGGATTCGACTCAGACAGTGACTCAGATTCAGA 2280
2280 TAGCGATTTCAGATTTCAGATAGCGGATTCAGATTCAGCGAGTGATTCGACTCAGACAGCGA 2339
2281 TAGCGATTTCAGATTTCAGATAGCGGATTCAGATTCAGCGAGTGATTCGACTCAGACAGCGA 2340
2340 TTCTGACTCCGACAGTGATTCGACTCAGACAGCGGATTCAGATTCGCGACAGTGATTCGGA 2399
2341 TTCTGACTCCGACAGTGATTCGACTCAGACAGCGGATTCAGATTCGCGACAGTGATTCGGA 2400
2400 CTCAGATAGCGATTCGACTCAGATAGCGGATTCAGATTCAGACAGCGGATTCAGATTCAGA 2459
2401 CTCAGATAGCGATTCGACTCAGATAGCGGATTCAGATTCAGACAGCGGATTCAGATTCAGA 2460
2460 CAGCGATTTCAGATTTCAGATAGCGGATTCAGATTCGCGACAGTGACTCAGATTCGCGACAGTGA 2519
2461 CAGCGATTTCAGATTTCAGATAGCGGATTCAGATTCGCGACAGTGACTCAGATTCGCGACAGTGA 2520
2520 CTCGCGATTTCAGATAGCGGATTCAGATTCGCGACAGTGACTCAGATTCGCGACAGTGACTCAGA 2579
2521 CTCGCGATTTCAGATAGCGGATTCAGATTCGCGACAGTGACTCAGATTCGCGACAGTGACTCAGA 2580
2580 CTCAGACAGTGATTCGCGATTTCAGCGAGTGATTCGCGATTTCAGATAGTGATTCGCGACTCCGA 2639
2581 CTCAGACAGTGATTCGCGATTTCAGCGAGTGATTCGCGATTTCAGATAGTGATTCGCGACTCCGA 2640
2640 CAGTGACTCGGATTCAGATAGCGGATTCAGATTCGCGATAGCGGATTCGCGATTTCAGATAGCGGA 2699
2641 CAGTGACTCGGATTCAGATAGCGGATTCAGATTCGCGATAGCGGATTCGCGATTTCAGATAGCGGA 2700
2700 TTCCGAGCTCAGATAGCGGATTCAGATTCAGACAGCGGATTCAGATTCAGACAGCGGATTCAGA 2759
2701 TTCCGAGCTCAGATAGCGGATTCAGATTCAGACAGCGGATTCAGATTCAGACAGCGGATTCAGA 2760
2760 TTCCAGACAGCGGATTCAGACAGTGACTCAGATTCAGATAGTGATTCGCGATTTCAGCGAGTGA 2819
2761 TTCCAGACAGCGGATTCAGACAGTGACTCAGATTCAGATAGTGATTCGCGATTTCAGCGAGTGA 2820
2820 TTCCAGACTCAGGTTAGTGACTCCGATTCATCAAGTGATTCGCGACTCAGAAAGTGATTCAAA 2879
```

```
Db 2821 TTCCAGACTCAGGTTAGTGACTCCGATTCATCAAGTGATTCGCGACTCAGAAAGTGATTCAAA 2880
QY 2880 TAGCGATTCCGAGTCCAGGTTCTAAACAATAATGATGTTCCGCTCAATTTACCTCAAAAATCG 2939
Db 2881 TAGCGATTCCGAGTCCAGGTTCTAAACAATAATGATGTTCCGCTCAATTTACCTCAAAAATCG 2940
QY 2940 TACTAATGCTTCTTAATAAATAATGAGCTTAAAGATAGTAAAGAACCAATTTACCAGATACAGG 2999
Db 2941 TACTAATGCTTCTTAATAAATAATGAGCTTAAAGATAGTAAAGAACCAATTTACCAGATACAGG 3000
QY 3000 TTCTGAGATGAGCAATATACGCTCAATATTTGGGATTTATTAGCATCAATAGGTTTCATT 3059
Db 3001 TTCTGAGATGAGCAATATACGCTCAATATTTGGGATTTATTAGCATCAATAGGTTTCATT 3060
QY 3060 ACTACTTTTTCAGAGAAAAGAAAATAAGATAGAAATAGTAAATAGTAAATAGTAAATAGTAA 3119
Db 3061 ACTACTTTTTCAGAGAAAAGAAAATAAGATAGAAATAGTAAATAGTAAATAGTAAATAGTAA 3120
QY 3120 TTAATCATATGATTCATGAAAGAACCCCTTAAAGAGTGCTTCTTTTACTTGGATTTTCC 3179
Db 3121 TTAATCATATGATTCATGAAAGAACCCCTTAAAGAGTGCTTCTTTTACTTGGATTTTCC 3180
QY 3180 AATATATTTGTTGTAATATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3239
Db 3181 AATATATTTGTTGTAATATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3240
QY 3240 GTAGATGTTTATATAATTTGGCTTGGCGAAAATAAGGTTGTAAGGTAGGTTGTTAATTAG 3299
Db 3241 GTAGATGTTTATATAATTTGGCTTGGCGAAAATAAGGTTGTAAGGTAGGTTGTTAATTAG 3300
QY 3300 GGAAATAATGAGGAGAAAATAACAGTTTGAAATAATAATTTGCTAGTTTATCATTTGGGAGCAT 3359
Db 3301 GGAAATAATGAGGAGAAAATAACAGTTTGAAATAATAATTTGCTAGTTTATCATTTGGGAGCAT 3360
QY 3360 TATGCTATACAAATTTGGGAAAAGTAATCGTCCGAGTCCAGTGGTTCTGGGAGGAAGA 3419
Db 3361 TATGCTATACAAATTTGGGAAAAGTAATCGTCCGAGTCCAGTGGTTCTGGGAGGAAGA 3420
QY 3420 ATCCATATCTATCTGAGTCTGTTGAAACTGACTAATAATAATAATAATAATAATAATAATAATA 3479
Db 3421 ATCCATATCTATCTGAGTCTGTTGAAACTGACTAATAATAATAATAATAATAATAATAATAATA 3480
QY 3480 AAGAGTATAAGAAAAGCTT 3498
Db 3481 AAGAGTATAAGAAAAGCTT 3499

RESULT 3
AAV74898
ID AAV74898 standard; DNA; 4709 BP.
XX
AC AAV74898;
XX
DT 16-MAR-1999 (first entry)
XX
Staphylococcus aureus contig SEQ ID #587.
DE
KW Computer readable medium; vaccine; S.aureus infection; immunodetection; cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy; skin infection; surgical wound infection; scalded skin syndrome; toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
Key Location/Qualifiers
FH misc_feature /tag= a
FT 1681..1740
FT /note="these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT misc_feature 3481..3540
FT /tag= b
```


/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

EP786519-A2.
30-JUL-1997.
07-JAN-1997; 97EP-00100117.
05-JAN-1996; 96US-0009861P.
(HUMA-) HUMAN GENOME SCI INC.
Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
MPI; 1997-374922/35.
Polynucleotide(s) and proteins derived from *Staphylococcus aureus* - stored on computer readable medium and used in the production of anti-*S. aureus* vaccines.
Claim 1; Page 1519-1521; 3271pp; English.
This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the *S. aureus* DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against *S. aureus* infection. The polypeptides can also be used in a kit for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the *S. aureus* DNA sequences contained on the computer readable medium
Sequence 4709 BP; 1657 A; 704 C; 908 G; 1313 T; 0 U; 127 Other;
Query Match 93.8%; Score 3279.6; DB 2; Length 4709;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 3346; Conservative 1; Mismatches 131; Indels 21; Gaps 4;
1 GGTACCATTAATTACATCTGCTTTTGAAGAAATATGATTTCAAGCTAGGATTACATTA 60
1204 GGTACCATTAATTACATCTGCTTTTGAAGAAATATGATTTCAAGCTAGGATTACATTA 1263
61 GGTAGAGTTCATATTAATAATAAATAATGTTTGCAATCAATCGTACGTTGCTGTTGTA 120
1264 GGTAGAGTTCATATTAATAATAAATAATGTTTGCAATCAATCGTACGTTGCTGTTGTA 1323
121 ATTCTTAATAATAGCAATAATAAATGTTTGTAGTAAGTATTAATTGTGATATAAAA 180
1324 ATTCTTAATAATAGCAATAATAAATGTTTGTAGTAAGTATTAATTGTGATATAAAA 1383
181 TATCATACAAATTAATTGCTATTAATGCAATTTTGTATGTAATTTCCATTAAACAGAGATT 240
1384 TATCATACAAATTAATTGCTATTAATGCAATTTTGTATGTAATTTCCATTAAACAGAGATT 1443
241 AATATATATC-TTAAAGGATATATAGTTTAATATAAATGACCTTTTAAAGAGGGAATAA 299
1444 AATATATATCTTTAAAGGATATATAGTTTAATATAAATGACCTTTTAAAGAGGGAATAA 1503
300 AATCAATATGAAGAAAAAGAAAAACACGCAATTTGGAAAAAATCGATTGGCGTTC 359
1504 AATGAATATGAAGAAAAAGAAAAACACGCAATTTGGAAAAAATCGATTGGCGTTC 1563

QY 360 AGTCCTTGTAGTAGTAACTCGTTTAACTCGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAG 419
DB 1564 AGTCCTTGTAGTAGTAACTCGTTTAACTCGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAG 1623
QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATTCAAG 479
DB 1624 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATTNNN 1683
QY 480 TAGCGTTAGTCTGCACCTAAACAGAGGACACAAAAGTGAAGTACTATAAAACATGCTC 539
DB 1684 NNN 1743
QY 540 AACACTAATAATGCGAAACGAGTGTGGCGCAAAATCCAGCACAAACAGGAAACGACACA 599
DB 1744 AAACACTAATAATGCGAAACGAGTGTGGCGCAAAATCCAGCACAAACAGGAAACGACACA 1803
QY 600 ATCATCATCAACAAATGCACTACCGAAGAAACCGCGTAACTGCTGAAGCTACTACTAC 659
DB 1804 ATCATCATCAACAAATGCACTACCGAAGAAACCGCGTAACTGCTGAAGCTACTACTAC 1863
QY 560 GACAAAGCAATCAAGCTTAATACACCGCAACAACTCAATCAAGCAATCAAAATGCGGAGGA 719
DB 1864 GACAAAGCAATCAAGCTTAATACACCGCAACAACTCAATCAAGCAATCAAAATGCGGAGGA 1923
QY 720 ATTAGTGAATCAACAAAGTAAATGAACGACCTTTTAAATGATACTAAATCAGTATCACTGT 779
DB 1924 ATTAGTGAATCAACAAAGTAAATGAACGACCTTTTAAATGATACTAAATCAGTATCACTGT 1983
QY 780 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACAAACGCAAGATCTTCAAC 839
DB 1984 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACAAACGCAAGATCTTCAAC 2043
QY 840 TGAAGCAACCTTCAACCAATGAATCAGCTCCACAGAGTACAGATCGCAATTAATAAGA 899
DB 2044 TGAAGCAACCTTCAACCAATGAATCAGCTCCACAGAGTACAGATCGCAATTAATAAGA 2103
QY 900 TGTAAGTAAATCAACGCGTAAATCAAGTGGCGCTAGAGTGAAGAGCATTAGTTAGCGGC 959
DB 2104 TGTAAGTAAATCAACGCGTAAATCAAGTGGCGCTAGAGTGAAGAGCATTAGTTAGCGGC 2163
QY 960 AGTAGCTGCAGATCAACCGGAGTGGCAGACAGATATTACGAATCAGTTGACGAATGTAC 1019
DB 2164 AGTAGCTGCAGATCAACCGGAGTGGCAGACAGATATTACGAATCAGTTGACGAATGTAC 2223
QY 1020 AGTTGGTATTGACTCTGTACGACTGTATCCGACCAACAGCAGTTATGTCAAACTGAA 1079
DB 2224 AGTTGGTATTGACTCTGTACGACTGTATCCGACCAACAGCAGTTATGTCAAACTGAA 2283
QY 1080 TTATGGTATTTCAGTGCTAAATCTGCTGTTTAAAGGTGACACATTCAAAATTAACGTACC 1139
DB 2284 TTATGGTATTTCAGTGCTAAATCTGCTGTTTAAAGGTGACACATTCAAAATTAACGTACC 2343
QY 1140 TAAAGAAATTAACCTTAATGCTGTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
DB 2344 TAAAGAAATTAACCTTAATGCTGTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2403
QY 1200 AGATCAAGTATTCGCAATGCTGTAATCGATGATGATGATGATGATGATGATGATGATGAT 1259
DB 2404 AGATCAAGTATTCGCAATGCTGTAATCGATGATGATGATGATGATGATGATGATGATGAT 2463
QY 1260 AGATCAAGTATTCGCAATGCTGTAATCGATGATGATGATGATGATGATGATGATGATGAT 1319
DB 2464 AGATCAAGTATTCGCAATGCTGTAATCGATGATGATGATGATGATGATGATGATGATGAT 2523
QY 1320 CCCTGAAAAATGTTAAAGAGCAGTAAATGTCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1379
DB 2524 CCCTGAAAAATGTTAAAGAGCAGTAAATGTCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2583
QY 1380 AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAAGTTTATTAACCTTATCTAT 1439
DB 2584 AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAAGTTTATTAACCTTATCTAT 2643

1440 TAAAGGTACAAATTGACCAAAATCGATAAATAAATAAATGATCGTACAGCAATTTATGT 1499
 2644 TAAAGGTACAAATTGACCAAAATCGATAAATAAATAAATGATCGTACAGCAATTTATGT 2703
 1500 CAAATCAAGTGGAGATAAAGTTATTCGGCGGGTTTAAACAGGTAATTTAAACCAATATC 1559
 2704 CAAATCAAGTGGAGATAAAGTTATTCGGCGGGTTTAAACAGGTAATTTAAACCAATATC 2763
 1560 GGATAGTAAATGATTAATAGATCAGCAAAATAAATAAATGATCGTACAGCAATTTATGT 1619
 2764 GGATAGTAAATGATTAATAGATCAGCAAAATAAATAAATGATCGTACAGCAATTTATGT 2823
 1620 TGCAGCTGATTTATCTGAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCCTAA 1679
 2824 TGCAGCTGATTTATCTGAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCCTAA 2883
 1680 TAGTGTGAATATTAATTCGCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 1739
 2884 TAGTGTGAATATTAATTCGCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 2943
 1740 TGATCAAAATTAACAACACCGTATATAGTGTGTTAATGCTCATATTTGATCGCAATAGCAA 1799
 2944 TGATCAAAATTAACAACACCGTATATAGTGTGTTAATGCTCATATTTGATCGCAATAGCAA 3003
 1800 AGGTGATTTAGCTTTTACGTTCAACTTTTATATGGGTATACTCGAATATATTTTGGCGCTC 1859
 3004 AGGTGATTTAGCTTTTACGTTCAACTTTTATATGGGTATACTCGAATATATTTTGGCGCTC 3063
 1860 TAGTGTGGAATTAATTCGCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 1919
 3064 TAGTGTGGAATTAATTCGCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 3123
 1920 TAAACCAAGTTTTCCTGAACCACTGATGAGCTGTGTAATTTGAACCAATTCAGGATCAGA 1979
 3124 TAAACCAAGTTTTCCTGAACCACTGATGAGCTGTGTAATTTGAACCAATTCAGGATCAGA 3183
 1980 TTCAGATTCGACCCAGGTTTCAGATTCGCGACGATTTCTAATTCAGATAGCGGTTTCTAGA 2039
 3184 TTCAGATTCGACCCAGGTTTCAGATTCGCGACGATTTCTAATTCAGATAGCGGTTTCTAGA 3243
 2040 TTCGGGTAGTGTATTCATATCAGATAGTGTTCAGATTCAGCGAGTGTTCAGATTCAGC 2099
 3244 TTCGGGTAGTGTATTCATATCAGATAGTGTTCAGATTCAGCGAGTGTTCAGATTCAGC 3303
 2100 AAGTGTATTCAGATTCAGCGAGTGTTCAGATTCAGCGAGTGTTCAGATTCAGCGAGGCA 2159
 3304 AAGTGTATTCAGATTCAGCGAGTGTTCAGATTCAGCGAGTGTTCAGATTCAGCGAGGCA 3363
 2160 TTCGACTCAGACAAATGACTTCGATTCAGATAGCGATTTCTGATCAGACAGTGTGACTCAGA 2219
 3364 TTCGACTCAGACAAATGACTTCGATTCAGATAGCGATTTCTGATCAGACAGTGTGACTCAGA 3423
 2220 TTCGACAGTGTGACTCAGATTCAGATAGCGATTTCTGATCAGACAGTGTGACTCAGATTCAGA 2279
 3424 TTCGACAGTGTGACTCAGATTCAGATAGCGATTTCTGATCAGACAGTGTGACTCAGATTCAG 3483
 2280 TAGCGATTCAGATTCAGATAGCGATTTCTGATCAGACAGTGTGACTCAGACAGCGCA 2339
 3484 NNN 3543
 2340 TTCGACTCCGACAGTGTATTCGACTCAGACAGCGATTCAGATTCGACAGTGTGACTTCGCA 2399
 3544 TTCGACTCCGACAGTGTATTCGACTCAGACAGCGATTCAGATTCGACAGTGTGACTTCGCA 3603
 2400 CTCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGA 2459
 3604 CTCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGA 3663
 2460 CAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGTGACTCAGATTCGACAGTGTG 2519
 3664 CAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGTGACTCAGATTCGACAGTGTG 3723
 2520 CTCGGATTCAGATAGCGATTCAGATTCGACAGTGTGACTCAGATTCGACAGTGTGACTCAGA 2579

3724 CTCGATTCAGATAGCGATTCAGATTCGACAGTGTGACTCAGATTCGACAGTGTGACTCAGA 3783
 2580 CTCGACAGTGTATTCGATTCAGCGAGTGTATTCGATTCAGATAGTGTGACTCAGATTCGCA 2639
 3784 CTCGACAGTGTATTCGATTCAGCGAGTGTATTCGATTCAGATAGTGTGACTCAGATTCGCA 3843
 2640 CAGTGTATTCGATTCAGATAGCGATTCAGATTCGACAGTGTGACTCAGATTCGATTCAGATAGCGA 2699
 3844 CAGTGTATTCGATTCAGATAGCGATTCAGATTCGACAGTGTGACTCAGATTCGATTCAGATAGCGA 3903
 2700 TTCGACTCAGATAGCGATTCAGATTCAGATTCGACAGTGTGACTCAGATTCGATTCAGATAGCGA 2759
 3904 TTCGACTCAGATTCAGATTCAGATTCAGATTCGACAGTGTGACTCAGATTCGATTCAGATAGCGA 3945
 2760 TTCGACAGCGATTCAGACAGTGTGACTCAGATTCAGATTCGATTCGATTCAGCGAGTGA 2819
 3946 TTCGACAGCGATTCAGACAGTGTGACTCAGATTCAGATTCGATTCGATTCAGCGAGTGA 4005
 2820 TTCGACTCAGGTGTGACTTCGATTCAGATTCAGATTCGATTCAGATTCGATTCAGATTCAGATTCAG 2879
 4006 TTCGACTCAGGTGTGACTTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 4065
 2880 TAGCGATTCGAGTGTGACTTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2939
 4066 TAGCGATTCGAGTGTGACTTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 4125
 2940 TACTAATGCTTCTAATAAATAAAGTGTGACTTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2999
 4126 TACTAATGCTTCTAATAAATAAAGTGTGACTTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 4185
 3000 TTCGAAAGTGTGACTTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 3059
 4186 TTCGAAAGTGTGACTTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 4245
 3060 ACTACTTTTCAG 3119
 4246 ACTACTTTTCAG 4305
 3120 TTAATCATATGATTCATGAG 3179
 4306 TTAATCATATGATTCATGAG 4364
 3180 AATATATTTTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3239
 4365 AATATATTTTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4424
 3240 GTAGATTTTATATAATTTTGGTTCGCGAG 3299
 4425 GTAGATTTTATATAATTTTGGTTCGCGAG 4484
 3300 GGAG 3359
 4485 GGAG 4544
 3360 TATGTGTATCAG 3419
 4545 TATGTGTATCAG 4604
 3420 ATCCATATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 3479
 4605 ATCCATATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 4663
 3480 AAGAGTATTAAG 3498
 4664 AAGAGTATTAAG 4682

RESULT 4
 ABT14896
 ID ABT14896 standard; DNA; 2799 BP.
 XX
 AC ABT14896;

XX 06-MAR-2003 (first entry)
DT Pathogen specific antigen related staphylococcal DNA SEQ ID No 41.
DE
XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
XX hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW autoimmune disease; HIV; hepatitis; gene; ds.
XX Staphylococcus sp.
OS
XX WO200259148-A2.
XX
XX 01-AUG-2002.
XX
XX 21-JAN-2002; 2002WO-EP000546.
XX
XX 26-JAN-2001; 2001AT-00000130.
XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX
XX Meinke A, Nagy E, Von Ahlsen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vytvytska O, Etz H, Dryia A, Weichhart T, Hafner M;
PI Tempelmaier B;
XX
XX WPI; 2003-075410/07.
XX
XX Identifying, isolating and producing hyperimmune serum-reactive antigens
XX from a pathogen, for preparing vaccine or medicament for treating or
XX preventing e.g. staphylococcal infections, comprises providing antibody
XX preparation.
XX
XX Example 7; Page 151; 252pp; English.
XX
XX The invention relates to a novel method for identifying, isolating and
XX producing hyperimmune serum-reactive antigens from a pathogen, tumour,
XX allergen, a tissue or host prone to auto-immunity, where the antigens are
XX used in a vaccine, comprises providing antibody preparation from a plasma
XX pool of a type of animal, or individual sera with antibodies against the
XX specific pathogen, tumour, allergen, tissue or host prone to auto-
XX immunity. The hyperimmune serum-reactive antigens comprising any of the
XX 62 sequences of 53-2261 amino acids fully defined in the specification,
XX or their hyperimmune fragments are useful for the manufacture of a
XX pharmaceutical preparation, particularly a vaccine against staphylococcal
XX infections or colonisation against *S. aureus* or *S. epidermidis*. The
XX preparation of antibodies is useful for the manufacture of a medicament
XX for treating or preventing staphylococcal infections or colonisation
XX against *S. aureus* or *S. epidermidis*. The antibody preparations may also
XX be used for diagnostic and imaging purposes. Other conditions that can be
XX treated include cancer, autoimmune diseases or infections caused by viral
XX (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
XX polynucleotide sequence represents staphylococcal DNA relating to the
XX method for identifying and producing pathogen specific antigens of the
XX invention
SQ Sequence 2799 BP; 981 A; 527 C; 582 G; 709 T; 0 U; 0 Other;

Query Match 80.0%; Score 2797.4; DB 8; Length 2799;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2798; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 301 ATGAATATGAAGAAAAAGAAAAACAGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA 360
DB 1 ATGAATATGAAGAAAAAGAAAAACAGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA 60

QY 361 GTGCTTAGTACGTTAAATCGTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAGT 420
DB 61 GTGCTTAGTACGTTAAATCGTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAGT 120

QY 421 GAAATATGTTACGCAATCTGATAGCGCAAGTAACGAACGCAAAAGTAATGATTCAGT 480
DB 121 GAAATATGTTACGCAATCTGATAGCGCAAGTAACGAACGCAAAAGTAATGATTCAGT 180

QY 481 AGCGTTAGTGTGCTGCACTTAAAAACAGACGACACAAACGTTAGTGATCTATAAACATCGTCA 540
DB 181 AGCGTTAGTGTGCTGCACTTAAAAACAGACGACACAAACGTTAGTGATCTATAAACATCGTCA 240

QY 541 AACACTAATATGCGGAAAAAGAGTGTGCGCAAAATCCAGCACACAGAGAAACGACACAA 600
DB 241 AACACTAATATGCGGAAAAAGAGTGTGCGCAAAATCCAGCACACAGAGAAACGACACAA 300

QY 601 TCATCATCAAAATGCAACTACGGAAGAAAAACGCGGTAACTGTTGAGCTACTACTACG 660
DB 301 TCATCATCAAAATGCAACTACGGAAGAAAAACGCGGTAACTGTTGAGCTACTACTACG 360

QY 661 AACACGAATCAAGCTAATACACCGGCAACCAACTCAATCAAGCAATACAAATGCGGAGGAA 720
DB 361 AACACGAATCAAGCTAATACACCGGCAACCAACTCAATCAAGCAATACAAATGCGGAGGAA 420

QY 721 TTAGTGAATCAAAACGAATTAATGAAACGACTTTTAAATGATACTAATAACAGTATCTGTGA 780
DB 421 TTAGTGAATCAAAACGAATTAATGAAACGACTTTTAAATGATACTAATAACAGTATCTGTGA 480

QY 781 AATTCACTCAAAATTTCTAAGATGCGGAAAAATGTTTCAACACGCAAGTACTTCAACT 840
DB 481 AATTCACTCAAAATTTCTAAGATGCGGAAAAATGTTTCAACACGCAAGTACTTCAACT 540

QY 841 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAAGAT 900
DB 541 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAAGAT 600

QY 901 GTAGTTAATCAAGCGGTTAATACAAAGTGGCGCTTGAATAGAGAGCAATTAGTTTAGCGGCA 960
DB 601 GTAGTTAATCAAGCGGTTAATACAAAGTGGCGCTTGAATAGAGAGCAATTAGTTTAGCGGCA 660

QY 961 GTAGCTCAGATGACCGGCGAGCTGGCACAGATATTACGATCAGTTGACGAATGTGACA 1020
DB 661 GTAGCTCAGATGACCGGCGAGCTGGCACAGATATTACGATCAGTTGACGAATGTGACA 720

QY 1021 GTTGGTATTGACTCTGCTGACACTGTGTATCCGCAACCAAGCAGGTATTGTCAAACTGAAT 1080
DB 721 GTTGGTATTGACTCTGCTGACACTGTGTATCCGCAACCAAGCAGGTATTGTCAAACTGAAT 780

QY 1081 TATGGTTTTTCAAGTGCCTAATCTGCTGTTAAAGGTGACACATTCAAAATAAATCTGTACCT 1140
DB 781 TATGGTTTTTCAAGTGCCTAATCTGCTGTTAAAGGTGACACATTCAAAATAAATCTGTACCT 840

QY 1141 AAAGAAATTAACCTTAAATGGTGTAACTTCACTGCTAAAGTGCCCACTTATGCGTGA 1200
DB 841 AAAGAAATTAACCTTAAATGGTGTAACTTCACTGCTAAAGTGCCCACTTATGCGTGA 900

QY 1201 GATCAAGTATTGGCAAAATGGTAAATCGATAGTGTATGTTTATATACATTATACA 1260
DB 901 GATCAAGTATTGGCAAAATGGTAAATCGATAGTGTATGTTTATATACATTATACA 960

QY 1261 GACTATGTAATATCTAAAGATGATGTAAGCAAACTTTGACCATGCGCGCTTATTTGAC 1320
DB 961 GACTATGTAATATCTAAAGATGATGTAAGCAAACTTTGACCATGCGCGCTTATTTGAC 1020

QY 1321 CCTGAAATGTTAAAAAGACAGAGTAATGTGACATTTGCTACTGTCATAGGTAGTACACA 1380
DB 1021 CCTGAAATGTTAAAAAGACAGAGTAATGTGACATTTGCTACTGTCATAGGTAGTACACA 1080

QY 1381 GCRAACAAAACAGTATTAGTAGATTATGAAAAAATATGTTAAAGTTTATACTTCTATT 1440
DB 1081 GCRAACAAAACAGTATTAGTAGATTATGAAAAAATATGTTAAAGTTTATACTTCTATT 1140

QY 1441 AAAGGTCAATTTGACCAAAATCGATAAAACAAATATACGTATCGTCAGACAAATTTATGTC 1500
DB 1141 AAAGGTCAATTTGACCAAAATCGATAAAACAAATATACGTATCGTCAGACAAATTTATGTC 1200

QY 1501 AATCCAAAGTGAGATAACGTTATTGCGCGGTTTAAACAGGTAAATTTAAACCAAAATACG 1560
DB 1201 AATCCAAAGTGAGATAACGTTATTGCGCGGTTTAAACAGGTAAATTTAAACCAAAATACG 1260

1561 GATAGTAATGCAATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAAT 1620
1261 GATAGTAATGCAATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAAT 1320
1621 GAGCTGATTTACTGAAAGTTACTTTGTAATCCAGAAAACTTTGAGGATGTCATAAT 1680
1321 GAGCTGATTTACTGAAAGTTACTTTGTAATCCAGAAAACTTTGAGGATGTCATAAT 1380
1681 AGTGCAATATTACATTTCCCAATCCAAATCAATATAAAGTATAGATTTAATAGCGCTGAT 1740
1381 AGTGCAATATTACATTTCCCAATCCAAATCAATATAAAGTATAGATTTAATAGCGCTGAT 1440
1741 GATCAAAATTACAAACCGGTATATAGTGTGTTAATGGTCAATATGATCCGATACAAA 1800
1441 GATCAAAATTACAAACCGGTATATAGTGTGTTAATGGTCAATATGATCCGATACAAA 1500
1801 GGTGATTTAGCTTCACTTTCACTTTATATGGGTATAAATCGAATATAATTTGGCGCTCT 1860
1501 GGTGATTTAGCTTCACTTTCACTTTATATGGGTATAAATCGAATATAATTTGGCGCTCT 1560
1861 ATGTGATGGGACAAACGAAGTATGATTTAATAGCGATCAGTTCTGTCGACGGTATCGAT 1920
1561 ATGTGATGGGACAAACGAAGTATGATTTAATAGCGATCAGTTCTGTCGACGGTATCGAT 1620
1921 AAACCAAGTTGTTCTCTGAACAACTGATGAGCCTGGTGAATTTGAACCAATTCAGAGGAT 1980
1621 AAACCAAGTTGTTCTCTGAACAACTGATGAGCCTGGTGAATTTGAACCAATTCAGAGGAT 1680
1981 TCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTCTAAATTCAGATAGCGGTT 2040
1681 TCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTCTAAATTCAGATAGCGGTT 1740
2041 TCGGGTAGTGATTCTACATCAGATAGTGTTCAGATTCAGCGAGTATTCAGATTCAGCA 2100
1741 TCGGGTAGTGATTCTACATCAGATAGTGTTCAGATTCAGCGAGTATTCAGATTCAGCA 1800
2101 AGTGATTCAGATTCAGCGAGTATTCAGATTCAGCGAGTATTCAGATTCAGCGAGGAT 2160
1801 AGTGATTCAGATTCAGCGAGTATTCAGATTCAGCGAGGATTCAGATTCAGCGAGGAT 1860
2161 TCGGATTCAGCAATGACTCGGATTCAGATTCAGATTCAGCGAGTATTCAGATTCAGCGAG 2220
1861 TCGGATTCAGCAATGACTCGGATTCAGATTCAGATTCAGCGAGTATTCAGATTCAGCGAG 2160
2221 TCGGATTCAGCAATGACTCGGATTCAGATTCAGATTCAGCGAGTATTCAGATTCAGCGAG 2280
1921 TCGGATTCAGCAATGACTCGGATTCAGATTCAGATTCAGCGAGTATTCAGATTCAGCGAG 1980
2281 AGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2340
1981 AGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2040
2341 TGTGATTCGACAGTATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2400
2041 TGTGATTCGACAGTATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2100
2401 TCAGATAGCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2460
2101 TCAGATAGCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2160
2461 AGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2520
2161 AGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2220
2521 TCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2580
2221 TCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2280
2581 TCAGACAGTATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2640
2281 TCAGACAGTATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2340
2641 AGTGACTCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2700

2341 AGTGACTCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2400
2701 TCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2760
2401 TCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2460
2761 TCAGACAGCGATTCAGACAGTACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2820
2461 TCAGACAGCGATTCAGACAGTACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2520
2821 TCAGACTCAGGTAGTACTCCGATTCATCAAGTATTCATCAAGTATTCATCAAGTATTCATCAAG 2880
2521 TCAGACTCAGGTAGTACTCCGATTCATCAAGTATTCATCAAGTATTCATCAAGTATTCATCAAG 2580
2881 AGCGATTCGAGTACTCAGATTCATCAAGTATTCATCAAGTATTCATCAAGTATTCATCAAGTAT 2940
2581 AGCGATTCGAGTACTCAGATTCATCAAGTATTCATCAAGTATTCATCAAGTATTCATCAAGTAT 2640
2941 ACTAATGCTTCTAATAAAAAATGAGGCTAAAGATAGTAAAGAACCATTCACAGATACAGGT 3000
2641 ACTAATGCTTCTAATAAAAAATGAGGCTAAAGATAGTAAAGAACCATTCACAGATACAGGT 2700
3001 TCTGAAGATCAAGCAAAATACGTCATTAATTTGGGGATTTAGTCATCAATAGTTCATTA 3060
2701 TCTGAAGATCAAGCAAAATACGTCATTAATTTGGGGATTTAGTCATCAATAGTTCATTA 2760
3061 CTACTTTTCAGAGAAAAAAGAAATTAAGATAGAGAA 3099
2761 CTACTTTTCAGAGAAAAAAGAAATTAAGATAGAGAA 2799

RESULT 5

ACF73781
ID ACF73781 standard; DNA; 2781 BP.

XX ACF73781;

XX 20-NOV-2003 (first entry)

XX Staphylococcus aureus DNA #1461.

XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target; gene; ds.

XX Staphylococcus aureus.

XX WO200294868-A2.

XX 28-NOV-2002.

XX 27-MAR-2002; 2002WO-IB002637.

XX 27-MAR-2001; 2001GB-00007661.

XX (CHIR-) CHIRON SPA.

XX Masignani V, Mora M, Scarselli M;

XX WPI; 2003-120786/11.

XX P-PSDB; ABW72221.

XX New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
S. aureus, e.g. sepsis.

XX Claim 6; SEQ ID NO 2921; 49pp; English.

XX The invention relates to novel genes and encoded proteins from

CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
preventing infection due to Staphylococcus bacteria, specifically an

CC infection caused by *S. aureus*. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel *S. aureus* genes of the invention
XX

SQ Sequence 2781 BP; 973 A; 522 C; 578 G; 708 T; 0 U; 0 Other;
Query Match 78.5%; Score 2746.6; DB 8; Length 2781;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2777; Conservative 0; Mismatches 4; Indels 18; Gaps 1;
Qy 301 ATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTGGCGTTCAC 360
Db 1 ATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTGGCGTTCAC 60
Qy 361 GTGCTTAGTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAGT 420
Db 61 GTGCTTAGTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAGT 120
Qy 421 GAAATATAGTTACGCAATCTGATAGCGCAAGTAACGAAGCAAAAGTAATGATTCAAGT 480
Db 121 GAAATATAGTTACGCAATCTGATAGCGCAAGTAACGAAGCAAAAGTAATGATTCAAGT 180
Qy 481 AGCGTTAGTGTGCACCTTAAACACAGACGACACAAACGTTAGTGTACTTAAACATCGTCA 540
Db 181 AGCGTTAGTGTGCACCTTAAACACAGACGACACAAACGTTAGTGTACTTAAACATCGTCA 240
Qy 541 AACACTAATATGCGGAAACGAGTGTGGCGAAATCCAGCACACAGGAAGACACACAA 600
Db 241 AACACTAATATGCGGAAACGAGTGTGGCGAAATCCAGCACACAGGAAGACACACAA 300
Qy 601 TCATCATCAACAAATGCAACTTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTACG 660
Db 301 TCATCATCAACAAATGCAACTTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTACG 360
Qy 661 ACACGCAATCAAGTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA 720
Db 361 ACACGCAATCAAGTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA 420
Qy 721 TTAGTGAAATCAACCAAGTAATGAACGACTTTTAAATGATCTAATACAGTATCATCTGTA 780
Db 421 TTAGTGAAATCAACCAAGTAATGAACGACTTTTAAATGATCTAATACAGTATCATCTGTA 480
Qy 781 AATTCACTCAAAATCTTACAAATGCGGAAATGTTTCAACAAACGAGATACATCAACT 840
Db 481 AATTCACTCAAAATCTTACAAATGCGGAAATGTTTCAACAAACGAGATACATCAACT 540
Qy 841 GAAGCAACACCTTCAACCAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT 900
Db 541 GAAGCAACACCTTCAACCAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT 600
Qy 901 GTAGTTAATCAAGCGTTTATACAGTCCGCTAGATGAGAGCATTTAGTTTACGGCA 960
Db 601 GTAGTTAATCAAGCGTTTATACAGTCCGCTAGATGAGAGCATTTAGTTTACGGCA 660
Qy 961 GTAGCTGAGATGACCGGAGCTGGCACAGATATTACGAATCAGTTGACCAATGTGACA 1020
Db 661 GTAGCTGAGATGACCGGAGCTGGCACAGATATTACGAATCAGTTGACCAATGTGACA 720
Qy 1021 GTTGGTATTGACTCTGGTACGATGTTATCCGCAACGAGGTTATGTCAAACTGAAT 1080
Db 721 GTTGGTATTGACTCTGGTACGATGTTATCCGCAACGAGGTTATGTCAAACTGAAT 780
Qy 1081 TATGGTTTTTCAGTGCCTTAATCTGCTGTTAAAGTGACACATTCAAAATACCTGTACCT 1140
Db 781 TATGGTTTTTCAGTGCCTTAATCTGCTGTTAAAGTGACACATTCAAAATACCTGTACCT 840
Qy 1141 AAAGAAATTAACCTTAAATGGTGAACCTTCAACTGCTAAAGTGCCACCAATTTATGGCTGGA 1200
Db 841 AAAGAAATTAACCTTAAATGGTGAACCTTCAACTGCTAAAGTGCCACCAATTTATGGCTGGA 900
Qy 1201 GATCAAGTATGGCAAAATGGTGTAAATCGATAGTGAATGTATTTATACATTTACA 1260

Db 901 GATCAAGTATTTGGCAAAATGGTGTAAATCGATAGTGAATGTATTTATACATTTACA 960
Qy 1261 GACTATCTAATACTAAGATGATGTAAGCAACTTTGACCATTGCGCTTATATTGAC 1320
Db 961 GACTATCTAATACTAAGATGATGTAAGCAACTTTGACCATTGCGCTTATATTGAC 1020
Qy 1321 CCTGAAAATGTTAAAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACA 1380
Db 1021 CCTGAAAATGTTAAAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACA 1080
Qy 1381 GCAAAACAAACAGTATTAGTAGATTGAAAAATATGTTAAGTATTAACTTCTATT 1440
Db 1081 GCAAAACAAACAGTATTAGTAGATTGAAAAATATGTTAAGTATTAACTTCTATT 1140
Qy 1441 AAAGGTACAATTGACCAAAATCGATAAAACAAATTAATACGTATCTGCAGCAATTTATGTC 1500
Db 1141 AAAGGTACAATTGACCAAAATCGATAAAACAAATTAATACGTATCTGCAGCAATTTATGTC 1200
Qy 1501 AATCCAAAGTGAGATAACGTTATTGCGCCGGTTTTTAACAGGTAATTTAAAAACCAATACG 1560
Db 1201 AATCCAAAGTGAGATAACGTTATTGCGCCGGTTTTTAACAGGTAATTTAAAAACCAATACG 1260
Qy 1561 GATAGTAATGCAATTAATAGATCAGCAAAATACAAGTATTAAGTATATAAAGTAGATAAT 1620
Db 1261 GATAGTAATGCAATTAATAGATCAGCAAAATACAAGTATTAAGTATATAAAGTAGATAAT 1320
Qy 1621 GCAGCTGATTATCTGAAAGTACTTTGTGAATCCAGAAACTTTGAGGATGTCACATAAT 1680
Db 1321 GCAGCTGATTATCTGAAAGTACTTTGTGAATCCAGAAACTTTGAGGATGTCACATAAT 1380
Qy 1681 AGTGTGAATATTACATTTCCCAAAATCCAAATCAATATAAAGTAGAGTTTAAATACCCCTGAT 1740
Db 1381 AGTGTGAATATTACATTTCCCAAAATCCAAATCAATATAAAGTAGAGTTTAAATACCCCTGAT 1440
Qy 1741 GATCAAAATTAACAACCGGTATATAGTAGTTGTTAATGTCATATTTGATCCGAATAGCAAA 1800
Db 1441 GATCAAAATTAACAACCGGTATATAGTAGTTGTTAATGTCATATTTGATCCGAATAGCAAA 1500
Qy 1801 GGTGATTTAGCTTTACGTTCAACTTTTATATGGGTATTAATCGAATATAATTTGGCGTCT 1860
Db 1501 GGTGATTTAGCTTTACGTTCAACTTTTATATGGGTATTAATCGAATATAATTTGGCGTCT 1560
Qy 1861 ATGTCTATGGGACAAACGAAGTAGCATTTTAAACCGGATCAGGTTCTGTGACGGTATCGAT 1920
Db 1561 ATGTCTATGGGACAAACGAAGTAGCATTTTAAACCGGATCAGGTTCTGTGACGGTATCGAT 1620
Qy 1921 AAACAGGTTGTTCTGTGAACACCTGATGAGCTGGTGAAATTTGAACCAATTTCCAGAGGAT 1980
Db 1621 AAACAGGTTGTTCTGTGAACACCTGATGAGCTGGTGAAATTTGAACCAATTTCCAGAGGAT 1680
Qy 1981 TCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTCTTAATTCAGATAGCGGTTTCAGAT 2040
Db 1581 TCAGATTTGACCCAGCGTTTCAGATTTCTGGCAGCGATTCTTAATTCAGATAGCGGTTTCAGAT 1740
Qy 2041 TCCGGTATGATTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA 2100
Db 1741 TCCGGTATGATTCTCATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA 1800
Qy 2101 AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCAACGAGTTCGGACTCAGCGAGCGAT 2160
Db 1801 AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCAACGAGTTCGGACTCAGCGAGCGAT 1860
Qy 2161 TCCGACTCAGACAAATGACTCGGATTTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGAT 2220
Db 1861 TCCGACTCAGACAAATGACTCGGATTTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGAT 1920
Qy 2221 TCCGACAGTGAATTCAGATTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGATTCAGAT 2280
Db 1921 TCCGACAGTGAATTCAGATTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGATTCAGAT 1980
Qy 2281 AGCGATTCAGATTCAGATAGCGATTTCAGATTCGACAGTGAATTCGACTCAGACAGCGAT 2340

QY 481 AGCGTTAGTGTGCTACCTAAAAACAGACGACACAAAACGTGAGTGATCTATAAAAACATCGTCA 540
DB 181 AGCGTTAGTGTGCTACCTAAAAACAGACGACACAAAACGTGAGTGATCTATAAAAACATCGTCA 240
QY 541 AACACTAATATGCGGAAACGAGTGTGGCGGAAATCCAGGACACAGGAAACGACAAA 600
DB 241 AACACTAATATGCGGAAACGAGTGTGGCGGAAATCCAGGACACAGGAAACGACAAA 300
QY 601 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAGCTACTACTACG 660
DB 301 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAGCTACTACTACG 360
QY 661 ACAACGAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGAA 720
DB 361 ACAACGAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGAA 420
QY 721 TTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATACCTAAACAGATACATCTGTA 780
DB 421 TTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATACCTAAACAGATACATCTGTA 480
QY 781 AATTCACCTCAAAATCTTACAATCGGAAATGTTTCAACAAACGAGATACCTCAACT 840
DB 481 AATTCACCTCAAAATCTTACAATCGGAAATGTTTCAACAAACGAGATACCTCAACT 540
QY 841 GAAGCAACACCTTCAAAACAAATGAGTCCACAGAGTACAGATGCAAGTAATAAGAT 900
DB 541 GAAGCAACACCTTCAAAACAAATGAGTCCACAGAGTACAGATGCAAGTAATAAGAT 600
QY 901 GTAGTTAATCAAGCGGTTAATAACAAGTGGCGCTAGANTGAGAGCATTTAGTTAGCGCA 960
DB 601 GTAGTTAATCAAGCGGTTAATAACAAGTGGCGCTAGANTGAGAGCATTTAGTTAGCGCA 660
QY 961 GTAGCTGAGATGACCGGACGCTGGCAACAGATATTTACGAATCAGTTGACGAATGTGACA 1020
DB 661 GTAGCTGAGATGACCGGACGCTGGCAACAGATATTTACGAATCAGTTGACGAATGTGAAA 720
QY 1021 GTTGTATGACTCTGGTAACTGTTGATCCGACCAAGCAGGTTATGTCAAACTGAAT 1080
DB 721 GTTGTATGACTCTGGTAACTGTTGATCCGACCAAGCAGGTTATGTCAAACTGAAT 780
QY 1081 TATGTTTTTTCAGTGCCTAAATCTGCTGTTAAAGTGACACATTCAAAATTAATCTGACT 1140
DB 781 TATGTTTTTTCAGTGCCTAAATCTGCTGTTAAAGTGACACATTCAAAATTAATCTGACT 840
QY 1141 AAAAGAAATTAACCTTAAATGGTGTAATCTTCAACTGCTAAAGTCCCAATTAATGGCTGA 1200
DB 841 AAAAGAAATTAACCTTAAATGGTGTAATCTTCAACTGCTAAAGTCCCAATTAATGGCTGA 900
QY 1201 GATCAAGTATGCGAAATGGTGTAATCGATAGTGAATGTTTATTTATACATTTACA 1260
DB 901 GATCAAGTATGCGAAATGGTGTAATCGATAGTGAATGTTTATTTATACATTTACA 960
QY 1261 GACTATGTAATTAAGTAAAGTAAAGCACTTTGACCATGCCGCTTATATTGAC 1320
DB 961 GACTATGTAATTAAGTAAAGTAAAGCACTTTGACCATGCCGCTTATATTGAC 1020
QY 1321 CCTGAAATGTTTAAAAACAGCTAATGTGCAATTTGGCTACTGGCATAGGTAGTACAAACA 1380
DB 1021 CCTGAAATGTTTAAAAACAGCTAATGTGCAATTTGGCTACTGGCATAGGTAGTACAAACT 1080
QY 1381 GCAACAAACAGTATTTAGTAGATTTATGAAATAATGTTTAACTTTTAACTTATCTATT 1440
DB 1081 GCTAGTAAAGCAGTATTTAATGACTATGAGAAATATGGAACAATTCGAATTTATCAATTT 1140
QY 1441 AAAGGTACAAATGACCAATCGATAAACAATAATACGTATCGTCAGCAATTTATGTC 1500
DB 1141 AAAGGTACAAATGACCAATCGATAAACAATAATACGTATCGTCAGCAATTTATGTC 1200
QY 1501 AATCCAAGTGAATTAAGTATTTGCGCGGTTTTTAAACAGGTAATTTAAAAACCAATAGC 1560
DB 1201 AATCCAAGTGAATTAAGTATTTGCGCGGTTTTTAAACAGGTAATTTAAATCTCTAATA 1260
QY 1561 GATAGTAAATGATTAATAGATCAGCAAAATACAAAGTATTAAGTATTAAGTAGATAAT 1620

DB 1261 AAGAGTAATGCGTTAATAGATGCAAAAAACACTGATATTAAGTTTATAGAGTCGATAAT 1320
QY 1621 GCAGCTGATTTATCTGAAAGTACTTTCTGTAATCCAGAAAACCTTTGAGGATCTCACTAAT 1680
DB 1321 GCTAATGATTTATCTGAAAGTATTTATGTAATCTTAGGATTTTGAAGATTTAATAAT 1380
QY 1681 AGTGTGAATATTTACATTCCTCAAAATCCAAATCAATATAAAGTAGAGTTTAAATCGGCTGAT 1740
DB 1381 CAAAGTTAGATTTTCAATTTCCAAATGCTAATCAATACAAAGTAGAATTTCTTACGGACGAT 1440
QY 1741 GATCAAAATTACAAACCGGTATATAGTGTGTTAATGCTCATATTTGATCCGAATAGCAAA 1800
DB 1441 GACCAAAATTACAAACCGGTATATAGTGTGTTAATGCTCAATTTGATCCTCTAGTACA 1500
QY 1801 GGTGATTTAGCTTTTACGTTCAACTTTATATGCTGATAAATCTGAATATTAATTTGGCGCTCT 1860
DB 1501 GGTGATTTAGCACTACGTTTCGACATTTTATGCTTATGATTTCTAAATTTTATATGGAGATCT 1560
QY 1861 ATGTCTATGGGACAAACGAAGTAGCATTTTAAATACGGATCAGGTTCTGGTGACGATTCGAT 1920
DB 1561 ATGTCTATGGGACAAACGAAGTAGCATTTTAAATACGGATCAGGTTCTGGTGACGATTCGAT 1620
QY 1921 AAACCAAGTTGTTCTTGAACCACTGATGAGCCTGGTGAAATTTGAACCAATTTCCAGAGAT 1980
DB 1621 AAACCAAGTTGTTCTTGAACCACTGATGAGCCTGGTGAAATTTGAACCAATTTCCAGAGAT 1680
QY 1981 TCAGATTTCTGACCCAGGTTTCAGATTTCTGACGAGGATTTCTAAATTCAGATAGCGGTTTCAGAT 2040
DB 1681 TCAGATTTCTGACCCAGGTTTCAGATTTCTGACGAGGATTTCTAAATTCAGATAGCGGTTTCAGAT 1740
QY 2041 TCGGTAGTGTATTTTACATTCAGATAGTGTGTTTCAGATTCAGCGAGTGTATTCAGATTCAGCA 2100
DB 1741 TCGGTAGTGTATTTTACATTCAGATAGTGTGTTTCAGATTCAGCGAGTGTATTCAGATTCAGCA 1800
QY 2101 AGTGTATTCAGATTCAGCGAGTGTATTCAGATTCAGCAAGGATTCGACTCAGCGAGCGAT 2160
DB 1801 AGTGTATTCAGATTCAGCGAGTGTATTCAGATTCAGCAAGGATTCAGATTCAGCAAGTGTAT 1860
QY 2161 TCCGACTCAGACAATGACTTCGGATTCAGATAGCGATTTCTGACTCAGACAGATTCAGATTCAGAT 2220
DB 1861 TCAGATTCAGCAAGTGTATTCAGATTCAGCAAGGATTCAGATTCAGCAAGTGTATTCAGATTCAGAT 1920
QY 2221 TCCGACTCAGATTCAGATTCAGATAGCGATTTCTGACTCAGACAGTGTACTCAGATTCAGATTCAGAT 2280
DB 1921 TCAGCAAGCGATTCAGATTCAGCGAGCGATTCAGATTCAGCGAGCGATTCAGATTCAGCG 1980
QY 2281 AGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGTATTCGACTCAGACAGCGAT 2340
DB 1981 AGTGTATTCGACTCAGCGAGCGATTCAGATTCAGATAGTGTACTCAGATTCGATAGCGAT 2040
QY 2341 TCTGACTCCGACAGTGTATTCGACTCAGACAGCGATTCAGATTCGACAGATTCGATTCGAC 2400
DB 2041 TCCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCGACTCAGACAGCGATTCGATTCGAC 2100
QY 2401 TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGAT 2460
DB 2101 TCAGACAGTGTACTCAGATTCGATAGCGATTCGACTCAGACAGTGTACTCAGATTCGATTCGAT 2160
QY 2461 AGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGTACTCAGATTCGACAGTGTAC 2520
DB 2161 AGCGATTCAGATTCAGACAGTGTATTCAGATTCAGATAGCGATTCAGATTCGACAGTGTAC 2220
QY 2521 TCGGATTCAGATAGCGATTCAGATTCGACAGTGTACTCAGATTCGACAGTGTACTCAGATTCGAC 2580
DB 2221 TCAGATTCAGACAGCGATTCAGATTCGATAGCGATTCAGATTCGACAGTGTACTCAGATTCAGAT 2280
QY 2581 TCAGACAGTGTATTCGATTCAGCGAGTGTATTCGATTCAGATTCAGATTCGATTCGATTCGAC 2640
DB 2281 TCCGATAGTGTACTCAGATTCGCGAGTGTATTCAGATTCAGATTCAGATTCGATTCGATTCGAT 2340
QY 2641 AGTGTATTCGATTCAGATAGCGATTCAGATTCGCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2700

Db 2341 AGTGACTCAGACTCAGACAGTGGATTTCAGATTCAGATAGTGAATTCAGACTCAGACAGCGAT 2400
Qy TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2760
Db TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2460
Qy TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2814
Db TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2520
Qy TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2874
Db TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2580
Qy TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2934
Db TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2640
Qy TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2994
Db TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2700
Qy TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 3054
Db TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2760
Qy TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 3102
Db TCGGACTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGAT 2808

RESULT 7

AAS54809

ID AAS54809 standard; DNA; 3066 BP.

XX AAS54809;

XX 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #1121.

XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

XX antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207727P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0269308P.

XX (BLIT-) ELITRA PHARM INC.

XX Haseelbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU36950.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 8446; 511pp; English.

XX

XX

XX

XX

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIFO at
CC ftp.wifo.int/pub/published_pct_sequences
XX

SQ Sequence 3066 BP; 1058 A; 598 C; 638 G; 772 T; 0 U; 0 Other;

Query Match 59.9%; Score 2096.8; DB 4; Length 3066;

Best Local Similarity 83.0%; Pred. No. 0;

Matches 2545; Conservative 0; Mismatches 257; Indels 264; Gaps 3;

Qy 301 ATGAATATGAGAAAAAAGAAAAACACGCAATTCGGAATAAATCGATTGGCGTGGCTTCA 360

Db 1 ATGAATATGAGAAAAAAGAAAAACACGCAATTCGGAATAAATCGATTGGCGTGGCTTCA 60

Qy 361 GTCCTTGTAGGTACGTTAATCGGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAGT 420

Db 61 GTCCTTGTAGGTACGTTAATCGGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAGT 120

Qy 421 GAAATAGTCTTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAAAGTATGATCAAGT 480

Db 121 GAAACAGTGTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAAAGTATGATCAAGT 180

Qy 481 AGCGTTAGTCTGCACCTAAACACAGACACACAAACGTTAGTGTATCAATAAATCGTCA 540

Db 181 AGGTTAATGCTGCACTAAACACAGACACACAAACGTTAGTGTATCAATAAATCGTCA 240

Qy 541 AACACTAATTAATGCGAAACACGAGTGTGGCGCAAAATCCAGCAACAGGAAACGACAA 600

Db 241 AACACTAATTAATGCGAAACACGAGTGTGGCGCAAAATCCAGCAACAGGAAACGACAA 300

Qy 601 TCATCATCAACAAATGCAACTACGGAAGAACCGCGTAACTCTGTGAGCTACTACTACG 660

Db 301 TCAGCATCAACAAATGCAACTACGGAAGAACCGCGTAACTCTGTGAGCTACTACTACG 360

Qy 661 ACAACGAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGGGAGAA 720

Db 361 GCAACGAGCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGGGAGAA 420

Qy 721 TTAGTGAATCAACAAAGTAAATGAACGACTTTTAATGATCTAATACAGTATCATCTGTA 780

Db 421 TTAGTGAATCAACAAAGTAAATGAACGACTTTTAATGATCTAATACAGTATCATCTGTA 480

Qy 781 AATTCACTCAAAATTTCTACAAATGCGGAAATGTTTCAACACGCAAGTACTTCAACT 840

Db 481 AATTCACTCAAAATTTCTACAAATGCGGAAATGTTTCAACACGCAAGTACTTCAACT 540

Qy 841 GAAGCAACACTTTCAACAAATGAATCAGCTCCACAGATGACAGTCAAGTCAATTAAGAT 900

Db 541 GAAGCAACACTTTCAACAAATGAATCAGCTCCACAGATGACAGTCAAGTCAATTAAGAT 600

Qy 901 GTAGTTAATCAAGCGGTTAATACAAGTGGCGCTAGAAATGAGAGCATTTAGTTAGCGCA 960

Db 601 GTAGTTAATCAAGCGGTTAATACAAGTGGCGCTAGAAATGAGAGCATTTAGTTAGCGCT 660

Qy 961 GTAGTGCAGATGCAACCGGCTGCGACAGATATTACGAATCAGTTGACGAATGTGACA 1020

Db 661 GTAGTGCAGATGCAACCGGCTGCGACAGATATTACGAATCAGTTGACGAATGTGAAA 720

1021 GTTGGTATTAGCTCTGGTAAGCTGTGTATCCGACCAAGCAGGTTTATGTCAAACTGAAT 1080
Db
721 GTTACTATTGACTCTGGTAGCACTGTGTATCCGACCAAGCAGGTTTATGTCAAACTGAAT 780
Qy
1081 TATGGTTTTAGTGCCTAAATCTGTGTAAAGGTGACAAATTCAAAATAAATCTGTACCT 1140
Db
781 TATGGTTTTAGTGCCTAAATCTGTGTAAAGGTGACAAATTCAAAATAAATCTGTACCT 840
Qy
1141 AAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTCCCACTAATATGCTGGA 1200
Db
841 AAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTCCCACTAATATGCTGGA 900
Qy
1201 GATCAAGTATTGGCAAAATGGTAAATCGATAGTGTAAATGTTATTTATACATTTACA 1260
Db
901 GATCAAGTATTGGCAAAATGGTAAATCGATAGTGTAAATGTTATTTATACATTTACA 960
Qy
1261 GACTATGTAAATACHTAAAGATGATGTAAAGCAACTTTGACCAATGCCCGCTTATATTGAC 1320
Db
961 GACTATGTGTACTAAAGAAATGTAAACAGCTAATATTACTATGCCAGCTTATATTGAC 1020
Qy
1321 CCTGAAATGTTAAAAAGACAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAACA 1380
Db
1021 CCTGAAATGTTAAAAAGACAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAACA 1080
Qy
1381 GCAACAAAACAGTATTAGTATGATTAAGAAAATATGTTAACTTTTATAACTTATCTATT 1440
Db
1081 GCTAGTAAGACAGTATTAACTGACTATGAGAAATATGGCAATTCCTAATTTTATCAATT 1140
Qy
1441 AAAGTACAAATGACCAATCGATAAACAAATTAATACGTATCGTCGACAAATTTATGTC 1500
Db
1141 AAAGTACAAATGACCAATCGATAAACAAATTAATACGTATCGTCGACAAATTTATGTC 1200
Qy
1501 AATCAAGTGGAGATAAGCTTATTGCGCGGTTTTAAACAGGTAAATTTAAACCAAATACG 1560
Db
1201 AATCAAGCGAGATAAGCTTATTGCTGTATTACTGCTTTAAACAGGTAAATTTAAATCTAATACA 1260
Qy
1561 GATAGTAATGCAATTAATAGATCAGCAAAATACAAGTATTTAAAGTATATAAAGTAGATAAT 1620
Db
1261 AAGAGTAATGCTTAATAGATCCAAAACAACTGTATTTAAAGTTTATAGAGTAGATAAT 1320
Qy
1621 SCAGCTGATTTATCTGAAGTTACTTTGTGATCCAGAAACCTTTGAGAGTGTCACTAAT 1680
Db
1321 GCTAATGATTTATCTGAAGTTATTTATGTGAATCCTAGCGATTTTGAAGATGTAACTAAT 1380
Qy
1681 AGTGTGAATATTACATCCCAATCCAAATCAATATAAAGTAGAGTTTAAACGCGCTGAT 1740
Db
1381 CAAGTTAGAAATTCATTTCCAAATCGTAATCAATACAAGTAGAATTTCTTAAGCAAGAT 1440
Qy
1741 GATCAAAATTACAACCGGTATATAGTGTGTTAATGGTCAATTTGATCCGAAATAGCAAA 1800
Db
1441 GATCAAAATTACAACCGGTATATAGTGTGTTAATGGTCAATTTGATCCGAAATAGCAAA 1500
Qy
1801 GGTGATTTAGCTTTTACGTTCACTTTATATGGGTATTAATCGAATATATTTGGCGCTCT 1860
Db
1501 GGTGATTTAGCTTTTACGTTCACTTTATGGTTAAGATTTCTAATTTTATATGGAGATCT 1560
Qy
1861 ATGTCATGGGACAACGAAGTAGCATTTAATACGGATCAGGTTCTGTGTGACGGTATCGAT 1920
Db
1561 ATGTCATGGGACAACGAAGTAGCATTTAATACGGATCAGGTTCTGTGTGACGGTATCGAT 1620
Qy
1921 AAACCAAGTTGTTCTGGAACAACCTGATCAGCGCTGTGAAATTTGAACCAATTTCCAGAGAT 1980
Db
1621 AAACCTGTGTTCTGGAACAACCTGATGAGCGCTGTGAAATTTGAACCAATTTCCAGAGAT 1680
Qy
1981 TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGAGCGATTTCTAATTCAGATAGCGGTTTCAGAT 2040
Db
1681 TCAGATTTCTGACCCAGGTTTCAGATAGTGTTCAGATTTCTAATTCAGATAGCGGTTTCAGAT 1740
Qy
2041 TCGGTAAGTGTATTACATCAGATAGTGTGTTTCAATTCAGAGGATGATTCAGATTTACGA 2100
Db
1741 TCGGTAAGTGTATTACATCAGATAGTGTGTTTCAATTCAGAGGATGATTCAGATTTACGA 1800
Qy
2101 AGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGATTCGACTCAGCGAGCGAT 2160

1801 AGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCAAGTGATTCAGATTCAGCAAGCGAT 1860
Qy
2161 TCCGACTCAGACAATAGCTCCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2198
Db
1861 TCAGATTCAGCGAGTGATTCAGATTCAGCAAGCGATTCAGATTCAGCAAGCGATTCAGATTCAG 1920
Qy
2199 ----- 2198
Db
1921 TCAGCAAGCGATTCAGATTCAGACTCAGACAGTGACTCAGACAGTGATTCAGATTCGACAGTGAC 1980
Qy
2199 -----TGACTCAGAC 2208
Db
1981 TCAGACTCAGATAGCGATTCAGATTCGATAGCGATTCGACTCAGACAGTGACTCAGAC 2040
Qy
2209 AGTGACTCAGATTCGACAGTGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2268
Db
2041 AGTGATTCAGATTCGACAGCGACTCAGATTCGATAGCGATTCGACTCAGATTCAGATTCAG 2100
Qy
2269 TCAGATTCAGATAGCGATTCAGATTCAGATTCAGATTCAGATTCGACAGTGATTCGAC 2328
Db
2101 TCAGATTCAGATAGCGATTCGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2160
Qy
2329 TCAGACAGCGATTCGACTCAGACAGTGATTCGACTCAGACAGTGATTCGACTCAGATTCGAC 2388
Db
2161 TCAGATAGCGATTCAGACTCAGACAGTGACTCAGATTCGATAGCGATTCGACTCAGATTCAG 2220
Qy
2389 AGTGATTCGACTCAGATAGCGATTCGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2448
Db
2221 AGTGATTCGATTCAGATAGCGATTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2280
Qy
2449 TCAGATTCAGACAGCGATTCAGATTCAGATTCAGATTCGACAGTGATTCGACAGTGATTCAG 2508
Db
2281 TCGGATTCAGATAGCGATTCAGATTCGATAGTGACTCAGACTCAGACAGCGATTCGAC 2340
Qy
2509 TCCGACAGTGACTCGGATTCAGATTCAGATTCAGATTCGACAGTGACTCAGATTCGAC 2568
Db
2341 TCAGACAGCGACTCGGATTCAGATTCAGATTCGATTCGATTCGACAGCGACTCAGACTCAG 2400
Qy
2569 AGTGACTCAGACTCAGACAGTGATTCGATTCGATTCAGCGAGTGATTCGATTCAGATTCAG 2628
Db
2401 AGCGATTCGAATTCGACAGTGATTCAGACTCAGACAGCGACTCAGATTCAGATTCAGATTCAG 2460
Qy
2629 TCCGACTCCGACAGTGACTCGGATTCAGATTCGATTCGACAGTGACTCAGATTCGAC 2688
Db
2461 TCCGACTCAGACAGTGACTCGGATTCAGATTCGATTCGATTCGACTCAGATTCAGATTCAG 2520
Qy
2689 TCAGATTCGATTCGACTCAGATTCAGATTCAGATTCAGATTCAGACAGCGATTCAGATTCAG 2748
Db
2521 TCAGATAGCGATTCAGATTCAGATTCGATTCGATTCGACAGCGACTCAGACTCAGATTCAG 2580
Qy
2749 AGCGATTCAGATTCAGACAGCGACTCAGAC----- 2778
Db
2581 AGTGACTCAGATTCAGATTCGATTCAGATTCAGATTCGATTCGATTCGACAGCGAC 2640
Qy
2779 -----AGTGACTCAGAT 2790
Db
2641 TCAGATTCAGCGAGTGATTCGACTCAGGTTAGTGACTCAGATTCAGATTCAGCGAGTGATTCGAC 2700
Qy
2791 TCAGATAGTGACTCGGATTCAGCGAGTGATTCAGATTCAGATTCAGGTTAGTGACTCAGATTCAG 2850
Db
2701 TCAGGTAAGTGACTCAGATTCAGCGAGTGATTCGACTCAGGTTAGTGACTCAGATTCAGATTCAG 2760
Qy
2851 AGTGAT----- 2856
Db
2761 AGTGATTCAGACTCAGACAGTGCTCTGATTCAGCGAGTGACACAGATTCAGACAGCGAC 2820
Qy
2857 TCCGACTCAGAAAAGTGATTCAGATTCAGGATTCGAGTTCCGAGTTCAGGTTCTAACAATATAGTGT 2916
Db
2821 TCTGACTCAGAAAAGTGATTCAGATTCAGGATTCGAGTTCCGACTCAGGTTCTAACAATATAGTGT 2880
Qy
2917 CGGCTAATTCACCTAAAAGTGATTCAGGTTAGTGTCTCTAATTAAGATGAGGCTAAGATAGT 2976

Db 2881 CGCGCTAAATTCACCTAAATAAGTCTAATGCTTCTAATAAAAAATGAGGCTAAAGATAGT 2940
QY 2977 AAGAACCATTACAGATACAGGTTCTGAAGATGAAGCAATACGTCACATAATTTGGGGA 3036
Db 2941 AAGAGCCATTACAGATACAGGTTCTGAAGGTAAGCGAATACGTCACATAATTTGGGGA 3000
QY 3037 TTATTAGCATCAATAGGTTTCATTACTACTTTTTCAGAGAAAAAAGAAAAATAAGATAAG 3096
Db 3001 TTATTAGCATCAATAGGTTTCATTACTACTTTTTCAGAGAAAAAAGAAAAATAAGATAAG 3060
QY 3097 AATAA 3102
Db 3061 AATAA 3066

RESULT 8

AAS51834

ID AAS51834 standard; DNA; 3063 BP.

XX

AC AAS51834;

XX

DT 13-FEB-2002 (first entry)

XX

DE Staphylococcus aureus DNA for cellular proliferation protein #251.

XX

KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.

XX

OS Staphylococcus aureus.

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-US009180.

XX

PR 21-MAR-2000; 2000US-0191078P.

XX

PR 23-MAY-2000; 2000US-0206848P.

XX

PR 26-MAY-2000; 2000US-0207272P.

XX

PR 23-OCT-2000; 2000US-0242578P.

XX

PR 27-NOV-2000; 2000US-0253625P.

XX

PR 22-DEC-2000; 2000US-0257931P.

XX

PR 16-FEB-2001; 2001US-0269308P.

XX

PA (ELI)- ELITRA PHARM INC.

XX

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX

XX MPI; 2001-61145570.

XX

XX P-PSDB; AAU33975.

XX

XX New polynucleotides for the identification and development of

XX PT antibiotics, comprise sequences of antisense nucleic acids.

XX

XX Claim 27; SEQ ID NO 4416; 511pp; English.

XX

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in

QY 1261 GACTATGTAATACTAAAGATGATGATAAGCAACTTTGACCATGCCCGCTTATATTGAC 1320
Db 961 GACTATGTTGATCTAAAGAAATGTAACAGCTAATATTACTATGCCAGCTTATATTGAC 1020
QY 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAACA 1380
Db 1021 CCTGAAATGTTTACAAGACAGGTAATGTGACATTTGCAACTGGCATAGGAACCAATACT 1080
QY 1381 GCAAAACAAAACAGTATTAGTAGATTGAAAAATATGTTAAGTTTATATAACTTACTATT 1440
Db 1081 GCTAGTAGACAGATTAATCGACTATGAGAAATATGACAATTCCTAATTAATCAATT 1140
QY 1441 AAGGTACAATGACCAATCGATAAAACAAATAATACGTATCGTCAAGCAAAATTTATGTC 1500
Db 1141 AAGGTACGATGATCAAAATCGATAAAACAAATAATACGTATCGTCAAGCAAAATTTATGTC 1200
QY 1501 AATCCAAAGTGAGATAAAGTTATTTCGGCGGGTTTTAAACAGGTAATTTTAAACCAATAGC 1560
Db 1201 AATCCAAAGCGAGATAAAGTTATTTCGGCTTTAAACAGGTAATTTTAAATTCCTAATACA 1260
QY 1561 GATAGTAATGCATTAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAAT 1620
Db 1261 AAGAGTAATGCTTAATAGATGCAAAACAACTGATTAATTAAGTTTATAGATAGATAAT 1320
QY 1621 GCAGCTGATTTATCTGAAAGTTACTTTGTGTAATCCAGAAACCTTTGAGGATGCTAAT 1680
Db 1321 GCTAATGATTTATCTGAAAGTTATTATGTGAATCTAGCGATTTTGAAGATGTAATAAT 1380
QY 1681 AGTGGAATATTACATTTCCCAATCCAAATCAATATAAGTAGATTTTAAATAGCGCTGAT 1740
Db 1381 CAAGTTAGAAATTTCAATTTCCAAATCTAATCAATACAAAGTAGAATTTTCCCTACGGACGAT 1440
QY 1741 GATCAAAATTAACAACCCGTATATAGTAGTTGTTAATGCTCATTTGATCCGAATAGCAAA 1800
Db 1441 GATCAAAATTAACAACCCGTATATTTAGTTGTTAATGCGCATTTGATCTCTAGACA 1500
QY 1801 GGTGATTTAGCTTTACGTTCAACTTTATATGGGTATAACTCGAATATAATTTGGCGCTCT 1860
Db 1501 GGTGATTTAGCACTACGTTTCGACATTTTATGGTTATGATTTCTAAATTTATATGGAGATCT 1560
QY 1861 ATGTCATGGGACAGGAAGTAGCATTTTAAACAGATCAGGTTCTGGTGACGGTATCGAT 1920
Db 1561 ATGTCATGGGACAGGAAGTAGCATTTTAAACAGATCAGGTTCTGGTGACGGTATCGAT 1620
QY 1921 AAACAGGTTGTTCTGGAACAACTGATGAGCCTGGTGAATTCGAACCAATTCACAGGAT 1980
Db 1621 AAACCTGTTGTTCTGGAACAACTGATGAGCCTGGTGAATTCGAACCAATTCACAGGAT 1680
QY 1981 TCAGATTCGACCCAGGTTTCAGATTCGCGACGATTCCTAATTCAGATAGCGGTTTCAGAT 2040
Db 1681 TCAGATTCGACCCAGGTTTCAGATAGTGGTTTCAGATTCCTAATTCAGATAGCGGTTTCAGAT 1740
QY 2041 TCGGTAAGTATCTTACATCAGATAGTGGTTTCAGATTCAGCGAGTTCAGATTCAGATTCAGCA 2100
Db 1741 TCGGTAAGTATCTTACATCAGATAGTGGTTTCAGATTCAGCGAGTTCAGATTCAGATTCAGCG 1800
QY 2101 AGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGATTCGAGCTCAGCGAGCGAT 2160
Db 1801 AGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCAAGCGATTCAGATTCAGATTCAGCAAGCGAT 1860
QY 2161 TCCGACTCAGCAATGATTCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2198
Db 1861 TCAGATTCAGCGAGTGATTCAGATTCAGCAAGCGATTCAGACTCAGCAAGCGATTCAGAC 1920
QY 2199 ----- 2198
Db 1921 TCAGCAAGCGATTCAGACTCAGACAGTACTCAGACAGTATTCAGATTCGACAGAGTAC 1980
QY 2199 ----- 2208
Db 1981 TCAGACTCAGATAGCGATTCAGATTCGATAGCGATTCGAGTTCAGACAGTACTCAGAC 2040
QY 2209 AGTGACTCAGATTCGACAGTACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2268

Db 2041 AGTGATTCAGATTCGACAGCGACTCAGATTCGATAGCGATTCGAGTTCAGATTCAGATTCAGAC 2100
QY 2269 TCAGATTCAGATAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTATTCGAGTTCGAC 2328
Db 2101 TCAGATTCAGATAGCGATTCGAGTTCGAGTTCAGATAGCGATTCAGATTCAGACAGTACTCGGAT 2160
QY 2329 TCAGACAGCGATTCGATTCGACAGCGATTCGAGTTCGAGTTCGACAGCGATTCAGATTCGAC 2388
Db 2161 TCAGATAGCGATTCAGACTCAGACAGTACTCGGATTCAGATAGCGATTCAGATTCAGACTCAGAC 2220
QY 2389 AGTGATTCGACTCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTC 2448
Db 2221 AGTGATTCGATTCAGATAGCGATTCGAGTTCGAGTTCAGATAGCGATTCAGATTCAGACAGTACT 2280
QY 2449 TCAGATTCAGACAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTACTCAGATTCAGATTC 2508
Db 2281 TCGGATTCAGATAGCGATTCAGATTCGAGTTCGAGTTCAGATTCAGACAGCGATTCGAGTTCGAG 2340
QY 2509 TCCGACAGTACTCGGATTCAGATAGCGATTCAGATTCGACAGTACTCAGATTCGAGTTCGAGTTCGAG 2568
Db 2341 TCAGACAGCGATTCGATTCAGATAGCGATTCGAGTTCGAGTTCGACAGCGATTCAGACTCAGAC 2400
QY 2569 AGTGACTCAGACTCAGACAGTATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2628
Db 2401 AGCGATTCGATTCGACAGTATTCAGACTCAGACAGCGATTCAGATTCAGATTCGAGTTCGAGTTC 2460
QY 2629 TCCGACTCCGACAGTACTCGGATTCAGATAGCGATTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTC 2688
Db 2461 TCCGACTCAGACAGTACTCGGATTCAGATAGCGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2520
QY 2689 TCAGATAGCGATTCGAGTTCAGATAGCGATTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2748
Db 2521 TCAGATAGCGATTCAGATTCAGATAGTATTCGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2580
QY 2749 AGCGATTCAGATTCAGACAGCGATTCAGAC----- 2778
Db 2581 AGTGACTCAGATTCAGATAGCGATTCAGACTCAGATAGTATTCGAGTTCGAGTTCGAGTTCGAGTTC 2640
QY 2779 -----AGTGACTCAGATTCAGATTCAGCGAGTATTCGAGTTCGAGTTCGAGTTCGAGTTC 2790
Db 2641 TCAGATTCAGCGAGTATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2700
QY 2791 TCAGATAGTACTCGGATTCAGCGAGTATTCAGATTCAGGTTAGTACTCCGATTCGATTCATCA 2850
Db 2701 TCAGGTAGTACTCAGATTCAGCGAGTATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCATCA 2760
QY 2851 AGTGATTCAGACTCAGACAGTTCGCTCTGATTCAGCGAGTACACAGAAATCAGAACGAC 2856
Db 2761 AGTGATTCAGACTCAGACAGTTCGCTCTGATTCAGCGAGTACACAGAAATCAGAACGAC 2820
QY 2857 TCCGACTCAGAAAGTATTCAGATAGCGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2916
Db 2821 TCTGACTCAGAAAGTATTCAGATAGCGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2880
QY 2917 CCGCTAATTCACCTAAATAGTATTAATGCTTCTAATAAAATGAGGCTAAAGATAGT 2976
Db 2881 CCGCTAATTCACCTAAATAGTATTAATGCTTCTAATAAAATGAGGCTAAAGATAGT 2940
QY 2977 AAAGAACCAATTCAGATAGCGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 3036
Db 2941 AAAGAGCCATTCAGATAGCGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 3000
QY 3037 TTATAGCATCAATAGGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3096
Db 3001 TTATAGCATCAATAGGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
QY 3097 AAA 3099
Db 3061 AAA 3063

ABK13616
 ID ABK13616 standard; DNA; 2792 BP.
 XX
 AC ABK13616;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE S. aureus gene for an antigenic protein #3.
 XX
 KW Antigenic protein; ds; gene; vaccine; SEREX; antibacterial;
 KW antiinflammatory; dermatological; antiulcer; tuberculosis;
 KW immunosuppressive; septicemia; food poisoning; skin disorders;
 KW peritonitis; endocarditis; tuberculosis; blood infection; sepsis;
 KW meningitis; pneumonia; stomach ulcer; gonorrhea; necrotising fasciitis;
 KW impetigo; Lyme's disease; gastro-enteritis; dysentery; shigellosis.
 XX
 OS Staphylococcus aureus
 XX
 XX NC0200198499-A1.
 XX
 XX 27-DEC-2001.
 XX
 XX 20-JUN-2001; 2001WO-GB0024685.
 XX
 XX 20-JUN-2000; 2000GB-00014907.
 XX
 XX (UYSH-) UNIV SHEFFIELD.
 XX (BIOS-) BIOSNEXUS INC.
 XX
 XX Foster S, McDowell P, Brummell K, Clarke S;
 XX
 XX NRI; 2002-106544/14.
 XX
 XX Identifying antigenic polypeptides expressed by pathogenic organisms
 PT e.g., Staphylococcus aureus during infection, by SEREX (serological
 PT identification of antigens by recombinant expression cloning) techniques.
 XX
 XX Claim 1; Page 33-34; 85pp; English.
 XX
 CC The invention relates to a method for identifying antigenic polypeptides
 CC expressed by pathogenic organisms e.g., Staphylococcus aureus during
 CC infection, by SEREX (serological identification of antigens by
 CC recombinant expression cloning) techniques. The method involves providing
 CC a nucleic acid library encoding genes/partial gene sequences of
 CC pathogenic organisms, transforming/transfecting the library into host
 CC cells, contacting the polypeptides expressed by the genes with autologous
 CC antisera derived from an animal infected with, or has been infected with
 CC the pathogen and purifying the nucleic acid encoding the polypeptide or
 CC partial polypeptide binding to the antisera. Also included are the
 CC nucleic acids and polypeptides isolated by the method, vectors and
 CC transformed cells expressing them, a vaccine comprising the polypeptide
 CC and the production of monoclonal antibodies against the polypeptides. The
 CC protein and vaccine are useful for immunising an animal (preferably
 CC human) against a pathogenic microbe. The proteins and antibodies are
 CC useful for manufacturing a medicament for treating Staphylococcus aureus-
 CC associated septicemia, food poisoning or skin disorders or
 CC Staphylococcus epidermidis-associated septicemia, peritonitis,
 CC endocarditis, tuberculosis, blood infections, sepsis, meningitis,
 CC pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, impetigo,
 CC Lyme's disease, gastro-enteritis, dysentery and shigellosis. The present
 CC sequence is an S. aureus gene isolated by the method of the invention
 CC which encodes an antigenic polypeptide
 XX
 SQ Sequence 2792 BP; 1078 A; 402 C; 525 G; 787 T; 0 U; 0 Other;
 Query Match 50.8%; Score 1776.4; DB 6; Length 2792;
 Best Local Similarity 99.9%; Pred. No. 5.5e-311;
 Matches 1788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 GGTACCATTAATACATCTGCTTTTGAAGGATATGATTTCAAGCTAGGATTACATTA 60
 DB 1003 GGTACCATTAATACATCTGCTTTTGAAGGATATGATTTCAAGCTAGGATTACATTA 1062

61 GGTAGAGTTCAVATTAAATAATAAAAAATGTTTGCATCAATCAATCGTAGCTGTGTTGTA 120
 1063 GGTAGAGTTCAVATTAAATAATAAAAAATGTTTGCATCAATCAATCGTAGCTGTGTTGTA 1122
 121 ATTCTTAAATAGCAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 180
 1123 ATTCTTAAATAGCAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1182
 181 TATCGATACAAATTAATTCGTATAATGCAATTTTAGTGTATAATTCATTAACAGAGATT 240
 1183 TATCGATACAAATTAATTCGTATAATGCAATTTTAGTGTATAATTCATTAACAGAGATT 1242
 241 AAATATATC-TTAAAGGGTATATAGTATAATAATAATAATAATAATAATAATAATAATA 299
 1243 AAATATATCTTTAAAGGGTATATAGTATAATAATAATAATAATAATAATAATAATAATA 1302
 300 AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTC 359
 1303 AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTC 1362
 360 AGTGCTTGTAGGTACGTTAATCGTTTTTGGACTACTCAGCAGTAAAGAACAGATGCAAG 419
 1363 AGTGCTTGTAGGTACGTTAATCGTTTTTGGACTACTCAGCAGTAAAGAACAGATGCAAG 1422
 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAAATGATCAAG 479
 1423 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAAATGATCAAG 1482
 480 TAGCGTTAGTGTGCACCTTAAACAGACGACACAAAGTGAGTGATCTAAACATCGTC 539
 1483 TAGCGTTAGTGTGCACCTTAAACAGACGACACAAAGTGAGTGATCTAAACATCGTC 1542
 540 ABAACATTAATAATGGCGAAGAGTGGCGGCAAAATCCAGCACACACAGAAACGACACA 599
 1543 ABAACATTAATAATGGCGAAGAGTGGCGGCAAAATCCAGCACACACAGAAACGACACA 1602
 600 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTAC 659
 1603 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTAC 1662
 660 GACAAAGAAATCAAGCTTAATACCGGCAACAACTCAATCAAGCAATCAAAATCGCGAGGA 719
 1663 GACAAAGAAATCAAGCTTAATACCGGCAACAACTCAATCAAGCAATCAAAATCGCGAGGA 1722
 720 ATTAGTGAATCAACAAAGTAAATGAACGACTTTTAAATGATACTAAATACAGTATCATCTGT 779
 1723 ATTAGTGAATCAACAAAGTAAATGAACGACTTTTAAATGATACTAAATACAGTATCATCTGT 1782
 780 AAATTTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACACGCAAGATACTTCAAC 839
 1783 AAATTTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACACGCAAGATACTTCAAC 1842
 840 TGAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATCAAGTATAAAGA 899
 1843 TGAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATCAAGTATAAAGA 1902
 900 TGTAGTTAATCAACGCGTTAATACAGTGGCGCTAGAAATGAGAGCAATTTAGTTTAGCGGC 959
 1903 TGTAGTTAATCAACGCGTTAATACAGTGGCGCTAGAAATGAGAGCAATTTAGTTTAGCGGC 1962
 960 AGTAGTGCAGATGACACCGGCGAGCTGGCAGAGATATTACGAATCAGTTGACGAATGTGAC 1019
 1963 AGTAGTGCAGATGACACCGGCGAGCTGGCAGAGATATTACGAATCAGTTGACGAATGTGAC 2022
 1020 AGTTGGTATTGACTCTGGTACGACTGTGATCCGACCAAGACGAGTTATGTCAAAACCTGAA 1079
 2023 AGTTGGTATTGACTCTGGTACGACTGTGATCCGACCAAGACGAGTTATGTCAAAACCTGAA 2082
 1080 TTATGGTTTTTTCAGTGGCTTAATTCGCTGTTAAAGGTGACACATTCAAAATACTGTACC 1139
 2083 TTATGGTTTTTTCAGTGGCTTAATTCGCTGTTAAAGGTGACACATTCAAAATACTGTACC 2142
 1140 TAAAGAATTAAACTTAAATGGTGTAACTTCAACTGCTAAAAGTGCCACCAATTTATGGCTGG 1199

Db 2143 TAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTGCCACCAATTAATGCTGCG 2202
Qy 1200 AGATCAAGTATTGGCAATGGTGAATCGATAGTGTGTAATGTTATTTATACATTAC 1259
Db 2203 AGATCAAGTATTGGCAATGGTGAATCGATAGTGTGTAATGTTATTTATACATTAC 2262
Qy 1260 AGACTATGTAAATACATAAAGATGATGTAAGCAACTTTGACCATGCCCGCTTATATTGA 1319
Db 2263 AGACTATGTAAATACATAAAGATGATGTAAGCAACTTTGACCATGCCCGCTTATATTGA 2322
Qy 1320 CCTGAAATGTTAAAGACAGGTAATGTGATCGTCTACTGGCATAGGTAGTACAAAC 1379
Db 2323 CCTGAAATGTTAAAGACAGGTAATGTGATCGTCTACTGGCATAGGTAGTACAAAC 2382
Qy 1380 AGCAACAACAAACAGTATTAGTATGATGTAAGCAATGTAAGCAATGTTTATACCTTATCTAT 1439
Db 2383 AGCAACAACAAACAGTATTAGTATGATGTAAGCAATGTAAGCAATGTTTATACCTTATCTAT 2442
Qy 1440 TAAAGGTACAAATGACCAAAATCGATAAATAAATAAATAAATAAATAAATAAATAAATAA 1499
Db 2443 TAAAGGTACAAATGACCAAAATCGATAAATAAATAAATAAATAAATAAATAAATAAATAA 2502
Qy 1500 CAATCAAGTGTAGATTAAGTTATGCGCCGGTTTTAAACAGGTAATTTAAACCAAAATAC 1559
Db 2503 CAATCAAGTGTAGATTAAGTTATGCGCCGGTTTTAAACAGGTAATTTAAACCAAAATAC 2562
Qy 1560 GGATAGTATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
Db 2563 GGATAGTATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2622
Qy 1620 TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAACTTTGAGGATGTCACATAA 1679
Db 2623 TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAACTTTGAGGATGTCACATAA 2682
Qy 1680 TAGTGTGAATATACATTCACCAATCCCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1739
Db 2683 TAGTGTGAATATACATTCACCAATCCCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2742
Qy 1740 TGATCAAAATTAACAACCGGTATATAGTATGATGATGATGATGATGATGATGATGATGATGAT 1789
Db 2743 TGATCAAAATTAACAACCGGTATATAGTATGATGATGATGATGATGATGATGATGATGATGAT 2792

RESULT 10

ID ADA89855 standard; DNA; 2792 BP.

XX AC ADA89855;

XX AC ADA89855;

DT 20-NOV-2003 (first entry)

XX Staphylococcus aureus antigenic partial gene sequence locus 83.

DE Staphylococcus aureus antigenic partial gene sequence locus 83.

XX Staphylococcus aureus antigenic partial gene sequence locus 83.

XX Staphylococcus aureus antigenic partial gene sequence locus 83.

KW antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;

KW antibacterial; neuroprotective; immunosuppressive; anti-inflammatory;

KW antitumor; immunostimulant; ophthalmological; pathogenic microbe;

KW bacteraemia; septic shock; organ infection; skin infection;

KW bacterial basal colonisation; bacterial eye infection; septicaemia;

KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;

KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;

KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;

KW gastro-enteritis; dysentery; shigellosis; skin disorder; gene; ds.

XX Staphylococcus aureus.

OS Staphylococcus aureus.

XX WO2003011899-A2.

XX 13-FEB-2003.

XX 02-AUG-2002; 2002WO-GB003606.

XX 02-AUG-2001; 2001GB-00018825.

XX 02-AUG-2001; 2001GB-00018825.

PR 09-JAN-2002; 2002GB-00000349.
XX (UYSH-) UNIV SHEFFIELD.
PA (BIOS-) BIOSYNEXUS INC.
XX Foster S, Mond J, Clarke S, McDowell P, Brummel K;
PI WPI; 2003-256434/25.
XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
PT useful as a vaccine for immunizing humans against e.g. bacteraemia, septic
PT shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
PT impetigo.
XX Claim 1; Page 104-105; 189pp; English.
XX The present invention describes an antigenic protein or its part, which
CC is for use as a vaccine. The antigenic protein is encoded by an isolated
CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
CC sequences (designated dnaSA and dna SE, respectively; and which encodes a
CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
CC composition comprising at least one antigenic protein; (2) a method of
CC immunising an animal against a disease or condition caused by a
CC pathogenic microbe by administering the antigenic protein or the vaccine;
CC (3) an antibody or its binding part obtainable by the method above; (4)
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
CC hybridoma cell line produced by the method of (4); and (6) identifying
CC opsonic antigens expressed by a pathogenic microbe. The antigenic
CC proteins have antibacterial, neuroprotective, immunosuppressive,
CC antiinflammatory, antitumor, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteraemia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
CC bacteraemia-associated food poisoning, blood infections, peritonitis,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
CC strep throat, streptococcal-associated toxic shock, necrotising
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
CC skin disorders, S. epidermidis-associated septicaemia, peritonitis or
CC endocarditis. The present sequence represents a S. aureus partial gene
CC sequence from the present invention.
XX Sequence 2792 BP; 1078 A; 402 C; 525 G; 787 T; 0 U; 0 Other;

Query Match 50.8%; Score 1776.4; DB 8; Length 2792;
Best Local Similarity 99.9%; Pred. NO. 5.5e-311;
Matches 1788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GGTACCAATAATTACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTACATTA 60

Db 1003 GGTACCAATAATTACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTACATTA 1062

Qy 61 GGTAGAGTTCATATTATAATAAAAAATGTTTGCATCAAAATCGTACGTTGCTGTTGTA 120

Db 1063 GGTAGAGTTCATATTATAATAAAAAATGTTTGCATCAAAATCGTACGTTGCTGTTGTA 1122

Qy 121 ATTCCTAAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTTATTTGGATATAATAA 180

Db 1123 ATTCCTAAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTTATTTGGATATAATAA 1182

Qy 181 TATCGATCAAAATTAATTTGCTATATGCAATTTTGTAGTATAAATCCATTACAGAGATT 240

Db 1183 TATCGATCAAAATTAATTTGCTATATGCAATTTTGTAGTATAAATCCATTACAGAGATT 1242

Qy 241 AAAATATATC-TTAAAGGATATATAGTTAATAATAAAATGACTTTTTAAAAAGAGGAATAA 299

Db 1243 AAAATATATC-TTAAAGGATATATAGTTAATAATAAAATGACTTTTTAAAAAGAGGAATAA 1302

Qy 300 AATGAATATGAGAAAGAAAGAAACACGCAATTCGGAAAAATCGATTGGCTGCTTC 359

Db 1303 AATGAATATGAGAAAGAAAGAAACACGCAATTCGGAAAAATCGATTGGCTGCTTC 1362


```
QY 360 AGTGTGTAGGTACGTTTAATCGGTTTTCGACTACTCAGCAGTAAGACGACATGCAAG 419
Db 1363 AGTGTGTAGGTACGTTTAATCGGTTTTCGACTACTCAGCAGTAAGACGACATGCAAG 1422
QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATCAAG 479
Db 1423 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATCAAG 1482
QY 480 TAGCGTTAGTGTGACCTTAATAACAGACGACACAAAGCGTGAGTGATCTAATAACATCGTC 539
Db 1483 TAGCGTTAGTGTGACCTTAATAACAGACGACACAAAGCGTGAGTGATCTAATAACATCGTC 1542
QY 540 AATACATTAATAAGCGGAACGAGTGTGGCGCAAAATCCAGACACACAGAGAAACGACACA 599
Db 1543 AATACATTAATAAGCGGAACGAGTGTGGCGCAAAATCCAGACACACAGAGAAACGACACA 1602
QY 600 ATCATCATCAAAATGCAACTACGGAAGAAACGCGGTAACCTGGTGAAGCTACTACTAC 659
Db 1603 ATCATCATCAAAATGCAACTACGGAAGAAACGCGGTAACCTGGTGAAGCTACTACTAC 1662
QY 660 GACAAAGATCAAGCTTAATACACGGCAACAACTCAATCAAGCAATACAAATCGGAGGA 719
Db 1663 GACAAAGATCAAGCTTAATACACGGCAACAACTCAATCAAGCAATACAAATCGGAGGA 1722
QY 720 ATTAGTGAATCAAAACAGTAATGAACGACTTTTATGATACTAATAACAGTATCATCTGT 779
Db 1723 ATTAGTGAATCAAAACAGTAATGAACGACTTTTATGATACTAATAACAGTATCATCTGT 1782
QY 780 AATTCACCTCAAAATTTCTCAAAATGCGGAAATGTTTCAACACGCAAGATACTTCAAC 839
Db 1783 AATTCACCTCAAAATTTCTCAAAATGCGGAAATGTTTCAACACGCAAGATACTTCAAC 1842
QY 840 TGAAGCAACACCTTCAAAACAGTAATGAACGACTCCACAGAGTACAGATCAAGTAATAAGA 899
Db 1843 TGAAGCAACACCTTCAAAACAGTAATGAACGACTCCACAGAGTACAGATCAAGTAATAAGA 1902
QY 900 TGTAGTTAATCAAGCGTTAATCAAGTGGCGCTAGAAATCAGAGACATTTAGTTTAGCGGC 959
Db 1903 TGTAGTTAATCAAGCGTTAATCAAGTGGCGCTAGAAATCAGAGACATTTAGTTTAGCGGC 1962
QY 960 AGTAGCTGCAGATGCACGGCAGCTGGCACAGATATTACGAATCAGTTGACGAATGTGAC 1019
Db 1963 AGTAGCTGCAGATGCACGGCAGCTGGCACAGATATTACGAATCAGTTGACGAATGTGAC 2022
QY 1020 AGTTGTTATGACTCTGCTGAGACTGTGTATCCGACCAAGCAGGTTATGTCACAACTGAA 1079
Db 2023 AGTTGTTATGACTCTGCTGAGACTGTGTATCCGACCAAGCAGGTTATGTCACAACTGAA 2082
QY 1080 TTATGTTTTCAGTGTGCTAATTTCTGCTGTTAAAGGTGACACATTCACAAATTAAGTACC 1139
Db 2083 TTATGTTTTCAGTGTGCTAATTTCTGCTGTTAAAGGTGACACATTCACAAATTAAGTACC 2142
QY 1140 TAAAGAAATTAACCTTAATGCTGTAACTTCAACTGTAAAGTGCACCAATTTAGGCTGG 1199
Db 2143 TAAAGAAATTAACCTTAATGCTGTAACTTCAACTGTAAAGTGCACCAATTTAGGCTGG 2202
QY 1200 AGATCAAGTATGCAATGCTGTAAATCGATAGTGGTAAGTTATTTATACATTTAC 1259
Db 2203 AGATCAAGTATGCAATGCTGTAAATCGATAGTGGTAAGTTATTTATACATTTAC 2262
QY 1260 AGACTATGTAATAACTAAGATGATGTAAGCAACCTTTGACCATGCCGCTTATATTGA 1319
Db 2263 AGACTATGTAATAACTAAGATGATGTAAGCAACCTTTGACCATGCCGCTTATATTGA 2322
QY 1320 CCCTGAAAATGTTAAAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGTAGTACAAC 1379
Db 2323 CCCTGAAAATGTTAAAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGTAGTACAAC 2382
QY 1380 AGCAACAAACAGTATTAGTAGATTATGAAAAATATGTTGAAGTTTATTAACCTTATCTAT 1439
Db 2383 AGCAACAAACAGTATTAGTAGATTATGAAAAATATGTTGAAGTTTATTAACCTTATCTAT 2442
```

```
QY 1440 TAAAGTACAATGACCAAAATCGATAAAACAAATAAATACGTATCGTCAGACAAATTTATGT 1499
Db 2443 TAAAGTACAATGACCAAAATCGATAAAACAAATAAATACGTATCGTCAGACAAATTTATGT 2502
QY 1500 CAATCCAAGTGGAGATAAAGCTTATTCGCGCGGTTTTTAAACAGGTAATTTAAAAACCAATAC 1559
Db 2503 CAATCCAAGTGGAGATAAAGCTTATTCGCGCGGTTTTTAAACAGGTAATTTAAAAACCAATAC 2562
QY 1560 GGNATGTAATGCAATTAATAGATCAGCAAAATACAACTATTAAGTATATAAAGTAGATAA 1619
Db 2563 GGNATGTAATGCAATTAATAGATCAGCAAAATACAACTATTAAGTATATAAAGTAGATAA 2622
QY 1620 TGCAGCTGATTTTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCATAA 1679
Db 2623 TGCAGCTGATTTTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCATAA 2682
QY 1680 TAGTGTGAATATTACATTTCCCAAAATCCAAATCAATATAAAGTAGAGTTTAATAGCGCTGA 1739
Db 2683 TAGTGTGAATATTACATTTCCCAAAATCCAAATCAATATAAAGTAGAGTTTAATAGCGCTGA 2742
QY 1740 TGATCAAAATTACAAACCGGTATATAGTAGTTGTTAATGGTCATATTGATC 1789
Db 2743 TGATCAAAATTACAAACCGGTATATAGTAGTTGTTAATGGTCATATTGATC 2792

RESULT 11
AAD46861
ID AAD46861 standard; DNA; 1560 BP.
XX
AC AAD46861;
XX
DT 27-JAN-2003 (first entry)
XX
DE Staphylococcus aureus Clf40 DNA.
XX
KW Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy;
KW gene; ds
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 1..1560
FT /tag= a
FT /product= "Staphylococcus aureus Clf40 protein"
FT /note= "CDS does not include start and stop codon"
XX
XX WO200272600-A2.
XX
XX 19-SEP-2002.
XX
XX 28-JAN-2002; 2002WO-US002296.
XX
XX 26-JAN-2001; 2001US-0264072P.
XX
XX 12-MAR-2001; 2001US-0274611P.
XX
XX 18-JUN-2001; 2001US-0298433P.
XX
XX 30-JUL-2001; 2001US-0308416P.
XX
XX (INHI-) INHIBITEX INC.
XX
XX Patti-JM Hutchins JT, Domanski P, Patel P, Hall A;
XX
XX WPI; 2002-759834/82.
XX
XX P-PSDB; AAE29262.
XX
XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for
XX treating or preventing Staphylococcus aureus infection e.g. wound
XX infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
XX a human or animal.
XX
XX Claim 10; Page 66-67; 80pp; English.
PS
```


XX The invention relates to monoclonal antibody which binds the clumping
CC factor A (ClfA) protein from *Staphylococcus aureus*. The anti-ClfA
CC monoclonal antibody is useful for treating or preventing *S. aureus*
CC infection in a human or animal, and for inhibiting the binding of
CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
CC *S. aureus* Clf40 protein, *S. aureus* Clf33 protein, or the *S. aureus* N3
CC protein is useful for inducing an immunological response in a human or
CC animal. These staphylococcal infections include wound infections, sepsis,
CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC present sequence is *Staphylococcus aureus* Clf40 DNA
XX
SQ Sequence 1560 BP; 584 A; 269 C; 299 G; 408 T; 0 U; 0 Other;
Query Match 44.6%; Score 1560; DB 6; Length 1560;
Best Local Similarity 100.0%; Pred. No. 5.5e-272; Indels 0; Gaps 0;
Matches 1560; Conservative 0; Mismatches 0;
418 AGTGAATAAGTGTTCACCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATCA 477
Db 1 AGTGAATAAGTGTTCACCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATCA 60
478 AGTAGCGTGTGCTGCACCTAAACAGACGACACAAAGCTGAGTGATCTAAACATCG 537
Db 61 AGTAGCGTGTGCTGCACCTAAACAGACGACACAAAGCTGAGTGATCTAAACATCG 120
538 TCAACACTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCACACAGCAAGAAACGACA 597
Db 121 TCAACACTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCACACAGCAAGAAACGACA 180
598 CAATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTCGTGAAGCTACTACT 657
Db 181 CAATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTCGTGAAGCTACTACT 240
658 ACGCAACGAATCAGCTAATACACCGGCAACAACTCAATCAGCAATCAAAATCGCGAG 717
Db 241 ACGCAACGAATCAGCTAATACACCGGCAACAACTCAATCAGCAATCAAAATCGCGAG 300
718 GAATAGTGAATCAACAAAGTAATCAAAACGACTTTTAATGATCTAATACAGTATCATCT 777
Db 301 GAATAGTGAATCAACAAAGTAATCAAAACGACTTTTAATGATCTAATACAGTATCATCT 360
778 GTAAATTCACCTCAAAATCTCAAAATCGGAAATGTTTCAACCAACCAAGATACCTCA 837
Db 361 GTAAATTCACCTCAAAATCTCAAAATCGGAAATGTTTCAACCAACCAAGATACCTCA 420
838 ACTGAGCAACACCTTCAAAACGAATCAATCAGCTCCACAGGTACAGATGCAAGTAATAA 897
Db 421 ACTGAGCAACACCTTCAAAACGAATCAATCAGCTCCACAGGTACAGATGCAAGTAATAA 480
898 GATGTAGTGAATCAACGCGTTAATAACAAGTGGCGCTAGAAATCAGAGCATTTAGTTAGCG 957
Db 481 GATGTAGTGAATCAACGCGTTAATAACAAGTGGCGCTAGAAATCAGAGCATTTAGTTAGCG 540
958 GCAGTAGCTGCAGATGCAACCGGACAGTGGCAAGATATACGAATCAGTTGACGAATG 1017
Db 541 GCAGTAGCTGCAGATGCAACCGGACAGTGGCAAGATATACGAATCAGTTGACGAATG 600
1018 ACAGTTGGTATGACTCTGTGACGACTGTATCGGACCAAGCAGGTTATGCAACTG 1077
Db 601 ACAGTTGGTATGACTCTGTGACGACTGTATCGGACCAAGCAGGTTATGCAACTG 660
1078 AATTATGTTTTTTCAGTGCCTAAATCTGCTGTAAAGGTGACACATTCAAATTAACCTGA 1137
Db 661 AATTATGTTTTTTCAGTGCCTAAATCTGCTGTAAAGGTGACACATTCAAATTAACCTGA 720
1138 CCTAAGAAATTAATTAATGGTGTAACTTCAACTGCTTAAGTGCCCAACCAATTAAGCT 1197
Db 721 CCTAAGAAATTAATTAATGGTGTAACTTCAACTGCTTAAGTGCCCAACCAATTAAGCT 780
1198 GGAGATCAAGTATGGCAAAATGGTGAATCATAGTGTGATGTTATGTTATTAACATTT 1257
Db 781 GGAGATCAAGTATGGCAAAATGGTGAATCATAGTGTGATGTTATGTTATTAACATTT 840

QY 1258 ACAGACTATGTAATACTAAAGATGATGTABAAGCAACTTTTGACCATGCCCGCTATATT 1317
Db 841 ACAGACTATGTAATACTAAAGATGATGTABAAGCAACTTTTGACCATGCCCGCTATATT 900
QY 1318 GACCCCTGAAAATGTTAAAAAGACAGGTAATGTGACATTGGCTACTGGCATAGGTAGTACA 1377
Db 901 GACCCCTGAAAATGTTAAAAAGACAGGTAATGTGACATTGGCTACTGGCATAGGTAGTACA 960
QY 1378 ACAGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAGTTTATTAACCTATCT 1437
Db 961 ACAGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAGTTTATTAACCTATCT 1020
QY 1438 ATTAAGGTACAAATGACCAATCGATAAACAATAATACGTATCGTCAGACAAATTTAT 1497
Db 1021 ATTAAGGTACAAATGACCAATCGATAAACAATAATACGTATCGTCAGACAAATTTAT 1080
QY 1498 GTCAATCCAAGTGGAGATAACGTTATTTCGCGCGGTTTTTAACAGGTAAATTTAAACCAAT 1557
Db 1081 GTCAATCCAAGTGGAGATAACGTTATTTCGCGCGGTTTTTAACAGGTAAATTTAAACCAAT 1140
QY 1558 ACGGATAGTAATGCAATTAATAGATCAGCAAAATACAGTATTAAAGTATATAAAGTAGAT 1617
Db 1141 ACGGATAGTAATGCAATTAATAGATCAGCAAAATACAGTATTAAAGTATATAAAGTAGAT 1200
QY 1618 AATGCAAGCTGATTATCTGAAAGTTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCATC 1677
Db 1201 AATGCAAGCTGATTATCTGAAAGTTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCATC 1260
QY 1678 AATAGTGTGAATATTACATTCCTCAATCCAAATCAATATAAAGTAGAGTTTAAATACGCT 1737
Db 1261 AATAGTGTGAATATTACATTCCTCAATCCAAATCAATATAAAGTAGAGTTTAAATACGCT 1320
QY 1738 GATGATCAAAATACAAACCGGTATATAGTAGTTGTTAATGTCATATTGATCCGAAATAGC 1797
Db 1321 GATGATCAAAATACAAACCGGTATATAGTAGTTGTTAATGTCATATTGATCCGAAATAGC 1380
QY 1798 AAAGTGATTATAGCTTTTACCTTCACTTTATATGGGTATACTCGAAATATAATTTGGCGC 1857
Db 1381 AAAGTGATTATAGCTTTTACCTTCACTTTATATGGGTATACTCGAAATATAATTTGGCGC 1440
QY 1858 TCTATGTCATGGGACAAACGAAAGTAGCAATTAATAACGGATCAGGTTCTGGTGACGGTATC 1917
Db 1441 TCTATGTCATGGGACAAACGAAAGTAGCAATTAATAACGGATCAGGTTCTGGTGACGGTATC 1500
QY 1918 GATAAACAGTGTGCTTCCGAAACCTGATGAGCTGGTGAATTTGAACCAATTTCCAGAG 1977
Db 1501 GATAAACAGTGTGCTTCCGAAACCTGATGAGCTGGTGAATTTGAACCAATTTCCAGAG 1560
RESULT 12
AAD46862
ID AAD46862 standard; DNA; 990 BP.
XX
XX AAD46862;
XX
XX 27-JAN-2003 (first entry)
XX
DE *Staphylococcus aureus* Clf33 DNA.
XX
XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein;
KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy;
KW gene; ds.
XX
OS *Staphylococcus aureus*.
XX
XX Key Location/Qualifiers
PH 1. .990
FT /*tag= a
FT /product= "Staphylococcus aureus Clf33 protein"
FT /transl_except= (pos:1..3, aa:Met-Val)
FT /note= "CDS does not include start and stop codon"

FT XX DN WO200272600-A2. /partial
PD XX 19-SEP-2002.
PF XX 28-JAN-2002; 2002WO-US002296.
PR XX 26-JAN-2001; 2001US-0264072P.
PR XX 12-MAR-2001; 2001US-0274611P.
PR XX 18-JUN-2001; 2001US-0298413P.
PR XX 30-JUL-2001; 2001US-0308116P.
XX XX (INHI-) INHIBITEX INC.
XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
XX WPI: 2002-759834/82.
XX P-PSDB: AAE29263.
XX New anti clumping factor A (ClfA) monoclonal antibody, useful for
XX treating or preventing Staphylococcus aureus infection e.g. wound
XX infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
XX a human or animal.
XX Claim 10; Page 70; 80pp; English.
XX The invention relates to monoclonal antibody which binds the clumping
XX factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
XX monoclonal antibody is useful for treating or preventing S. aureus
XX infection in a human or animal, and for inhibiting the binding of
XX staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
XX S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3
XX protein is useful for inducing an immunological response in a human or
XX animal. These staphylococcal infections include wound infections, sepsis,
XX impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
XX present sequence is Staphylococcus aureus Clf33 DNA
XX
SQ Sequence 990 BP; 352 A; 153 C; 192 G; 293 T; 0 U; 0 Other;
Query Match 28.3%; Score 990; DB 6; Length 990;
Best Local Similarity 100.0%; Pred. No. 3.4e-169;
Matches 990; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 961 GTAGCTCAGATGACACCGGACGCTGGACAGATATTACGATTCAGTTCAGCAATGTCACA 1020
DB 1 GTAGCTCAGATGACACCGGACGCTGGACAGATATTACGATTCAGTTCAGCAATGTCACA 60
QY 1021 GTTGGTATTGACTCTGGTACGACTGTGTATCCGACCAAGCAGGTATGTCAAACCTGAAT 1080
DB 61 GTTGGTATTGACTCTGGTACGACTGTGTATCCGACCAAGCAGGTATGTCAAACCTGAAT 120
QY 1081 TATGGTTTTTCAGTGCCTTAATTCCTGCTGTTAAAGGTGACACATTCAAATAAAGTGTACCT 1140
DB 121 TATGGTTTTTCAGTGCCTTAATTCCTGCTGTTAAAGGTGACACATTCAAATAAAGTGTACCT 180
QY 1141 AAGAATAAATTAATTAATGTTGTAATCTCACTGCTAAAGTGCACCAATTAATGGCTGGA 1200
DB 181 AAGAATAAATTAATTAATGTTGTAATCTCACTGCTAAAGTGCACCAATTAATGGCTGGA 240
QY 1201 GATCAAGTATTGGCAATGGTGTAAATCGATAGTGTATGTTATTTATACATTTACA 1260
DB 241 GATCAAGTATTGGCAATGGTGTAAATCGATAGTGTATGTTATTTATACATTTACA 300
QY 1261 GACTATGTAATACTAAAGATGATGTAAAGCAACTTTGACCATGCGCGCTTATATTGAC 1320
DB 301 GACTATGTAATACTAAAGATGATGTAAAGCAACTTTGACCATGCGCGCTTATATTGAC 360
QY 1321 CCTGAAATGTTAAAGACAGGTAAATGTGACATTTGGCTACTGTCATAGGTAGTACAACA 1380
DB 361 CCTGAAATGTTAAAGACAGGTAAATGTGACATTTGGCTACTGTCATAGGTAGTACAACA 420
QY 1381 GCAACAAACAGTATTAGTAGATTATGAAAAATATGTTAAAGTATTATTAACCTTATCTATT 1440

DB 421 GCAACAAACAGTATTAGTAGATTATGAAAAATATGTTAAAGTATTATTAACCTTATCTATT 480
QY 1441 AAAGGTACAATTGACCAATCGATATAAACAATAATACGTTATCGTCAGACAATTTATGTC 1500
DB 481 AAAGGTACAATTGACCAATCGATATAAACAATAATACGTTATCGTCAGACAATTTATGTC 540
QY 1501 AATCCAAGTGGAGATAACGTTATTGGCCCGGTTTTAACAGGTAAATTTAAACCAAAATACG 1560
DB 541 AATCCAAGTGGAGATAACGTTATTGGCCCGGTTTTAACAGGTAAATTTAAACCAAAATACG 600
QY 1561 GATAGTAATGCATTAAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAAT 1620
DB 601 GATAGTAATGCATTAAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAAT 660
QY 1621 GCAGCTGATTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCACCTAAT 1680
DB 661 GCAGCTGATTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCACCTAAT 720
QY 1681 AGTGTGAATATTACATTCCTCCAAATCCAAATCAATATAAAGTAGAGTTTAATACGCTGAT 1740
DB 721 AGTGTGAATATTACATTCCTCCAAATCCAAATCAATATAAAGTAGAGTTTAATACGCTGAT 780
QY 1741 GATCAAAATTACAACACCGTATATAGTAGTTGTTAAAGGTCAATATGATCCGAATAGCAAA 1800
DB 781 GATCAAAATTACAACACCGTATATAGTAGTTGTTAAAGGTCAATATGATCCGAATAGCAAA 840
QY 1801 GGTGATTAGCTTTACGTTCACTTTATATATGGGTATTAACCTGATATAATTTGGCGCTCT 1860
DB 841 GGTGATTAGCTTTACGTTCACTTTATATATGGGTATTAACCTGATATAATTTGGCGCTCT 900
QY 1861 ATGTCATGGGACAAACGAAGTAGCATTTAATAACGATCAGGTTCTGTCGACGGTATCGAT 1920
DB 901 ATGTCATGGGACAAACGAAGTAGCATTTAATAACGATCAGGTTCTGTCGACGGTATCGAT 960
QY 1921 AAACCAAGTTCCTCTGAAACCAACCTGATGAG 1950
DB 961 AAACCAAGTTCCTCTGAAACCAACCTGATGAG 990
RESULT 13
AAZ93533
ID AAZ93533 standard; DNA; 5406 BP.
XX AC AAZ93533;
XX DT 24-JUL-2000 (first entry)
XX DE Cell wall protein SdrF coding sequence.
XX KW SdrF; SdrH; SdrI; coagulase negative; staphylococcus; septicemia;
XX KW osteomyelitis; endocarditis; immune response; vaccine; graft; stent;
XX KW intravenous catheter; heart valve; cardiac; ss.
XX OS Staphylococcus sp.
FH Key Location/Qualifiers
FT CDS 1..5406
FT FT /*tag= a
FT FT /label= SdrF cell wall protein
FT FT 34..36
FT FT /*tag= b
FT FT /transl_except= TAA stop codon
FT FT 82..84
FT FT /*tag= c
FT FT /transl_except= TAA stop codon
FT FT 5311..5313
FT FT /*tag= d
FT FT /transl_except= TAG stop codon
FT FT 5320..5322
FT FT /*tag= e
FT FT /transl_except= TGA stop codon
FT FT 5377..5379

```

FT      /tag= f
FT      /transl_except= TGA stop codon
FT      5398..5400
FT      /tag= g
FT      /transl_except= TGA stop codon
XX
XX      WO200012689-A1.
XX      09-MAR-2000.
XX      31-AUG-1999; 99WO-US019728.
XX      31-AUG-1998; 98US-0098443P.
XX      25-JAN-1999; 99US-0117119P.
XX      (QUE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX      (TEXA ) UNIV TEXAS A & M SYSTEM.
XX      Poster TJ, Hook M, Davis S, Hartford O, McCreia K, Ni Eidhin D;
XX      P-PSDB; AAY83170.
XX      WPI; 2000-256637/22.
XX      DR      P-PSDB; AAY83170.
XX      CC      Recombinant or synthetic proteins from coagulase-negative staphylococci
XX      PT      useful for prevention, treatment and diagnosis of staphylococcal
XX      PT      infections bind soluble and immobilized fibrinogen.
XX      PS      Claim 3; Fig 2; 104pp; English.
XX      CC      Isolated staphylococcus-sdr cell wall proteins which bind both soluble
XX      CC      and immobilized fibrinogen are useful for treating or preventing
XX      CC      coagulase-negative staphylococcal infection such as septicemia,
XX      CC      osteomyelitis or endocarditis, and for inducing immune responses in
XX      CC      patients. The cell wall proteins are also useful for reducing coagulase-
XX      CC      negative staphylococci infection of indwelling medical devices such as
XX      CC      vascular grafts, vascular stents, intravenous catheters, artificial heart
XX      CC      valves and cardiac assist devices. The cell wall associated proteins are
XX      CC      able to inhibit staphylococcal adhesion to immobilised extracellular
XX      CC      matrix or host cells present on the surface of implanted biomaterials
XX      SQ      Sequence 5406 BP; 1982 A; 955 C; 1138 G; 1331 T; 0 U; 0 Other;
XX
XX      Query Match      17.7%; Score 619.2; DB 3; Length 5406;
XX      Best Local Similarity 78.8%; Pred. No. 3.5e-102;
XX      Matches 738; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
XX
XX      QY      1978 GATTCAGATCTGACCCAGGTTTCAGATTCGCGAGGATTCCTAATTCAGATAGCGGTTC 2037
XX      DB      4246 GACTCAGACGAGATAGTGTATTCGATTCGCGATTCGATAGTGTATTCGATAGTGTATTC 4305
XX      QY      2038 GATTCGGGTAGTGTATTCATCATCAGATAGTGTGTTCAGATTCAGCGAGTGTTCAGATTC 2097
XX      DB      4306 GACTCAGACAGTGTATTCGATTCGATAGCGACTCGGATTCAGATAGTGTATTCGACGCA 4365
XX      QY      2098 GCAAGTGTATTCAGATTCGCGAGTGTATTCAGATTCAGCAAGCGATTCGCGATTCAGCGAGC 2157
XX      DB      4366 GACAGTGTATTCAGATTCAGATAGTGTATTCGCGATTCGCGATTCGCGAGCAGCAGC 4425
XX      QY      2158 GATTCGCGACTCAGCAAGTGTATTCGATTCAGATAGCGATTCGATTCAGATAGTGTATTC 2217
XX      DB      4426 GATTCGCGACTCAGATAGTGTATTCAGATAGCGATTCGATTCGATAGTGTATTCGATAGC 4485
XX      QY      2218 GATTCGCGACTCAGATAGTGTATTCAGATAGCGATTCGATTCAGATAGTGTATTCGATAGC 2277
XX      DB      4486 GATGCAGACAGCGAGTTCGATTCGATAGTGTATTCGATAGTGTATTCGATAGTGTATTCG 4545
XX      QY      2278 GATACCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGTATTCGCGATTCAGACAGC 2337
XX      DB      4546 GATAGTGTATTCGATTCGATAGTGTATTCGCGAGTGTATTCGCGATTCGCGATTCGATAGC 4605
XX      QY      2338 GATTCGCGACTCAGATAGTGTATTCGCGATTCAGATAGCGATTCGATTCGCGAGTGTATTC 2397
XX      DB      4606 GATTCAGACTCCGACAGCGATTCAGATTCAGACAGCGACTCAGATTCGCGATAGTGTATTC 4665

```

```

QY      2398 GACTCAGATAGCGATTCGCGATTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTC 2457
DB      4666 GATTCAGACAGTGTATTCGATTCGATAGTGTATTCGATAGTGTATTCGATAGTGTATTC 4725
QY      2458 GACAGCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGTATTCGAGATTCGCGAGT 2517
DB      4726 GATAGCGACTCAGATTCAGACAGTGTATTCGCGACTCAGATAGTGTATTCGCGATTCAGACAGT 4785
QY      2518 GACTCGGATTCAGATAGCGATTCAGATTCGCGAGTGTATTCGCGAGTGTATTCGCGAGTGTATTC 2577
DB      4786 GATTCGCGATTCGATAGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCG 4845
QY      2578 GACTCAGACAGTGTATTCGCGATTCAGCGAGTGTATTCGCGATTCAGATAGTGTATTCGCGACTCC 2637
DB      4846 GACTCAGACAGCGACTTCGCGATTCAGATAGTGTATTCGCGACTCAGACAGCGATTCGCGATTC 4905
QY      2638 GACAGTGTATTCGCGATTCAGATAGCGACTCAGATTCGCGAGTGTATTCGCGATTCGCGATTC 2697
DB      4906 GATAGTGTATTCGCGATTCAGACAGTGTATTCGCGACTCAGACAGCGACTTCGCGATTCAGATAGT 4965
QY      2698 GATTCGCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTC 2757
DB      4966 GATTCGCGAGCGACAGCGACTTCGCGATTCAGATAGTGTATTCGCGAGCGACAGCGATTC 5025
QY      2758 GATTCAGACAGCGACTCAGACAGTGTATTCAGATTCAGATAGTGTATTCGCGATTCAGCGAGT 2817
DB      5026 GATAGTGTATTCGCGATTCAGACAGTGTATTCGCGACTCAGACAGCGATTCGCGATTCAGACAGT 5085
QY      2818 GATTCAGACTCAGGTAGTGTATTCGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2877
DB      5086 GACTCGGACTCAGATAGCGACTTCGCGATTCAGACAGTGTATTCGCGACTCAGATAGTGTATTC 5145
QY      2878 AATAGCGATTCGCGACTCAGGTTCCTAACATATGTA 2913
DB      5146 GATTCAGACAGCGACTTCGCGATTCGATAGTGTATTCGATAGTGTATTCGATAGTGTATTC 5181
XX
XX      RESULT 14
XX      AAZ51201
XX      ID      AAZ51201 standard; DNA; 5406 BP.
XX      AC      AAZ51201;
XX      DT      06-JUN-2000 (first entry)
XX      DE      Staph. epidermidis serine-aspartate repeat region protein SdrF gene.
XX      KW      Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAMM;
XX      KW      microbial surface components recognising adhesive matrix molecules;
XX      KW      collagen binding protein; CBP; CNA; fibrinogen binding protein;
XX      KW      Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;
XX      KW      fibronectin binding protein; Staphylococcus infection;
XX      KW      serine-aspartate repeat region protein; SDR protein; SdrF; ds.
XX      OS      Staphylococcus epidermidis.
XX      FH      Key
XX      CDS      Location/Qualifiers
XX              1..5406
XX              /*tag= a
XX              /product= "SdrF protein"
XX              /transl_except= (pos:34..36, aa:Xaa)
XX              /transl_except= (pos:82..84, aa:Xaa)
XX              /transl_except= (pos:5311..5313, aa:Xaa)
XX              /transl_except= (pos:5320..5322, aa:Xaa)
XX              /transl_except= (pos:5377..5379, aa:Xaa)
XX              /transl_except= (pos:5398..5400, aa:Xaa)
XX              /note= "The coding region does not include stop codon.
XX              xaa corresponds to in-frame stop codon"
XX              /partial
XX
XX      WO200012131-A1.

```

PD 09-MAR-2000.
 XX
 PF 31-AUG-1999; 99WO-US019727.
 XX
 XX 31-AUG-1998; 98US-0098439P.
 XX
 XX (INH-) INHIBITEX INC.
 PA (TEXA.) UNIV TEXAS A & M SYSTEM
 PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
 XX
 PI Patti JM, Foster TJ, Hook M;
 XX WPI; 2000-237781/20.
 XX P-PSDB; AAY70119.
 XX
 XX Composition used for generating immune response or for inhibiting
 PT microbial colonization in an animal comprises antibodies that bind
 PT collagen binding protein, fibrinogen binding protein and, optionally,
 PT fibronectin binding protein.
 XX
 XX Disclosure; Fig 3; 115pp; English.
 XX
 XX The patent discloses multicomponent vaccines containing selected
 CC combinations of bacterial binding proteins termed MSCRAMM (microbial
 CC surface components recognising adhesive matrix molecules) or their
 CC antibodies. A vaccine composition is provided that includes collagen
 CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein
 CC preferably Clumping factor A (ClfA) or Clumping factor B (ClfB), and
 CC optionally a fibronectin binding protein e.g. FnBP-A. The vaccines are
 CC useful for imparting protection against a broad spectrum of
 CC Staphylococcal strains and for inhibiting microbial colonisation,
 CC especially of Staphylococcus aureus, in an animal. The combinations can
 CC also be used to select donor blood pools for the preparation of purified
 CC blood products for passive immunisation. The present sequence is a gene
 CC encoding serine-aspartate repeat region protein, SdrF from Staphylococcus
 CC epidermidis. The Sdr protein is useful in vaccine preparation in
 CC combination with specific bacterial binding proteins. These vaccines can
 CC be used to treat a broad spectrum of bacterial infections, including
 CC those arising from both coagulase-positive and coagulase-negative
 CC bacteria
 XX
 XX
 SQ Sequence 5406 BP; 1982 A; 955 C; 1138 G; 1331 T; 0 U; 0 Other;
 Query Match 17.7%; Score 619.2; DB 3; Length 5406;
 Best Local Similarity 78.8%; Pred. No. 3.5e-102;
 Matches 738; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 1978 GATTTCAGATTCTGACCCAGGTTTCAGATTCTGGCAGCGGATTCTAATTCAGATAGCGGTCA 2037
 |||||
 DB 4246 GACTCAGACGACATAGTAGTATCCGATTCCGATAGCGGATTTCAGATTCTGATAGTACTCA 4305
 |||||
 QY 2038 GATTCGGGTAGTGTCTACATCAGATAGTAGTGGTTTCAGATTTCAGGAGTGTAGTATCA 2097
 |||||
 DB 4306 GACTCAGACAGTACTCAGATTCCGATAGCGACTCCGGATTTCAGATAGTGTATCCGACGCA 4365
 |||||
 QY 2098 GCAGGTGATTTCAGACTCAGCGAGTGTATTCAGATTTCAGCAGCGGATTCCGACTCAGCGAGC 2157
 |||||
 DB 4366 GACAGTACTCAGACTCAGATAGTAGTCTCGATTCCGATAGTGTATTCGACGCGAGCAGC 4425
 |||||
 QY 2158 GATTCGAGTTCAGACATGACTCGGATTTCAGATAGCGGATTCTGACTCAGACAGTGTACTCA 2217
 |||||
 DB 4426 GATTCGACTCAGATAGTACTCAGACGCGAGATAGTGTATTCGATTCCGATAGCGGATTTCG 4485
 |||||
 QY 2218 GATTCGACAGTGTACTCAGATTTCAGATAGCGGATTCTGACTCAGACAGTGTACTCAGATTCA 2277
 |||||
 DB 4486 GATGCACAGCGACTCCGATTTCAGATAGTGTATTCGACGCGAGACAGTGTACTCAGACTCA 4545
 |||||
 QY 2278 GATAGCGATTTCAGATTTCAGATAGCGGATTTCAGATTCCGACAGTGTATTCGACTCAGACAGC 2337
 |||||
 DB 4546 GATAGTACTCCGATTTCGATAGTGTATTCGACGCGAGACAGTGTATTCGATTTCGATAGC 4605
 |||||
 QY 2338 GATTCTGACTCCGACAGTGTATTCGACTCAGACAGCGGATTTCAGATTCCGACAGTGTATTC 2397
 |||||

RESULT 15

AAH54581

ID AAH54581 standard; DNA; 4627 BP.

XX

AC AAH54581;

XX

DT 03-SEP-2001 (first entry)

XX

DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3945.

XX

KW Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
 endocarditis; ds.

XX

QS Staphylococcus epidermidis.

XX

PN WO200134809-A2.

XX

PD 17-MAY-2001.

XX

HF 09-NOV-2000; 2000WO-US030782.

XX

PR 09-NOV-1999; 99US-0164258P.

XX

PA (GLAX) GLAXO GROUP LTD.

XX

PI Zimmerly WJ;

XX

DR WPI; 2001-316495/33.

XX

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 useful for vaccinating against infections, e.g. endocarditis.

XX

PS Claim 8; Page 1606-1608; 2188pp; English.

XX

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC *epidermidis* polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed *S. epidermidis* genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

XX
SQ Sequence 4627 BP; 1480 A; 849 C; 1077 G; 1221 T; 0 U; 0 Other;

Query Match 16.9%; Score 589.6; DB 4; Length 4627;

Best Local Similarity 71.9%; Pred. No. 7.4e-97;

Matches 817; Conservative 0; Mismatches 304; Indels 15; Gaps 3;

QY 1978 GATTTCAGATTTCTGACCCAGGTTTCAGATTTCTGCGAGCGATTTCTTAATTCAGATAGCGGTTCA 2037

DB 3454 GACTCGGATTCAGATAGTGTTCGGATTCAGACAGCGACTCGGATTCGACGATGATTCG 3513

QY 2038 GATTCGGGTAGTGATTTACATCAGATAGTGGTTCAGATTCAGCGAGTGATTCAGATTC 2097

DB 3514 GACGACAGAGTGACTCAGATTCAGACAGTGTTCGGACGCGACAGCGACTCGGATTC 3573

QY 2098 GCAAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCAGCGAGTTCGACTCAGCGAGC 2157

DB 3574 GATAGTGATTCGGACGCGACAGAGTGACTCAGATTCAGACAGTGACTCGGATTCAGACAGC 3633

QY 2158 -----GATTCGGACTCAGACAACTGACTCGGATTCAGATAGCGATTCGACTCAGACAGT 2211

DB 3634 GATTCGGATTCGATTCAGACAGTGACTCGGATTCAGACAGTGACTCAGACTCCGACAGT 3693

QY 2212 GACTCAGATTCGACAGTGACTCAGATTCAGATAGCGATTCGACTCAGACAGTGACTCA 2271

DB 3694 GATTCGGATTCAGATAGCGACTCAGACGACAGATAGTGATTCGATTCAGATAGCGACTCC 3753

QY 2272 GATTCAGATAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGATTCGAGCTCA 2331

DB 3754 GACGACAGATAGTGATTCGGACGCGACAGAGTGACTCAGATTCATACAGTGACTCAGATTC 3813

QY 2332 GACAGCGATTCGACTCGGACAGTGATTCGCGACTCAGACAGCGATTCAGATTCGACAGAGT 2391

DB 3814 GACAGTGATTCGGACGCGACAGTGACTCCGACTCCGACAGCGATTCAGACTCAGATAGT 3873

QY 2392 GATTCGCGACTCAGATAGCGATTCGCGACTCAGATAGCGACTCAGATTCAGACAGCGATTC 2451

DB 3874 GACTCAGACGCGACAGTGACTCGGACTCAGATAGTGATTCAGATTCGACAGCGATTC 3933

QY 2452 GATTCAGACAGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACAGTGACTCAGATTC 2511

DB 3934 GACTCAGATAGCGATTCGCGATTCAGACAGCGACTCCGACTCAGACAGTGATTCGCGATTC 3993

QY 2512 GACAGTGACTCGGATTCAGATAGCGATTCAGATTCGCGACAGTGACTCAGATTCGCGACAGT 2571

DB 3994 GACAGCGATTCGCGACTCAGATAGTGACTCAGACGCGAGATAGTGATTCGATTCAGATAGC 4053

QY 2572 GACTCAGACTCAGACAGTGATTCGGATTCAGCGAGTGATTCGGATTCAGATAGTGATTC 2631

DB 4054 GACTCCGATTCGATAGTGACTCCGATTCAGATAGCGACTCCGATTCAGATAGTGATTCG 4113

QY 2632 GACTCCGACAGTGACTCGGATTCAGATAGCGACTCAGATTCGATAGCGACTCCGATTC 2691

DB 4114 GACGACAGAGTGACTCGTACTCAGATAGTGACTCGGATTCAGACAGTGATTCGCGATTC 4173

QY 2692 GATAGCGGATTCGGACTCAGATAGCGATTCAGAATCAGACAGCGATTCAGAATCAGACAGC 2751

DB 4174 GATAGCGGATTCGGATTCGGATAGTGACTCGGATTCAGACAGTGATTCGCGACTCAGACAGC 4233

QY 2752 GATTCAGATTCAGACAG-----CGACTCAGACAGTGACTCAGATTCAGATAGTGACTCG 2805

DB 4234 GACTCCGATTCAGATAGTGATTCGCGACTCAGACAGCGATTCGCGATTCGATAGTGACTCG 4293

QY 2806 GATTCAGCGAGTGATTCAGACTCAGGTAGTGACTCCGATTCATCAAGTGATTCGCGACTCA 2865

DB 4294 GATTCAGACAGTGATTCGGACTCAGACAGCGATTCGCGATTCGCGATAGTGACTCGGATTC 4353

QY 2866 GAAAGTGATTCAAAATPAGCGATTCGCGAGTCAGGTTCTTAACAATAATGTAGTTCGCGCTAAT 2925

DB 4354 GACGTGATTCGGGCTCAGACAGCGATTCGCGATTCAGACAGTGACTCCGACTCAGATAGT 4413

QY 2926 TCACCTAAAAATGGTACTTAATGCTTCTTAATAAAAAATGAGGCTAAAGATAGTAAAGAACCA 2985

DB 4414 GACTCCGATTCAGACAGCGACTCGGATTCGTATAAAAATGCAAAAAGA---TAAATTACCT 4470

QY 2986 TTACCAGATACAGGTTCTGAAGATGAACCAATAGCTCACTAATTTGGGGATTATTAGCA 3045

DB 4471 GATACAGGAGCAAAATGAAGATCATGATTCCTAAGGCGACATTAATTGGAACTTTATTTCGA 4530

QY 3046 TCAATAGGTTTCATTACTACTTTTCAGAGAAGAAAAAGAAAAATAAGATAAGAAAAATA 3101

DB 4531 GGTTTAGGAGCATTTATTATTAGGAAGAGCTCGTAAAAAAGATATAAGAAAAATA 4586

Search completed: November 5, 2004, 06:04:58

Job time : 1579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2004, 05:38:03 ; Search time 276 Seconds
(without alignments)
9008.476 Million cell updates/sec

Title: US-09-679-643-1

Perfect score: 3498

Sequence: 1 ggtaccataataacacatc.....gaagagtataaagaagctt 3498

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCFUS COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3498	100.0	3498	3	US-08-293-728-1
2	3498	100.0	3498	3	US-09-421-868-1
3	3399.6	97.2	4709	4	US-08-956-171E-587
4	3399.6	97.2	4709	4	US-08-956-171E-587
5	619.2	17.7	5406	4	US-08-781-986A-587
6	619.2	17.7	5406	4	US-08-781-986A-587
7	589.6	16.9	4627	4	US-09-386-959-1
8	584.8	16.7	1485	4	US-09-710-279-35
9	568.6	16.3	2969	4	US-09-200-650E-2
10	510	14.6	11050	4	US-08-956-171E-96
11	510	14.6	11050	4	US-08-781-986A-96
12	462	13.2	3600	4	US-09-147-405B-14
13	448.4	12.8	4358	4	US-08-956-171E-454
14	448.4	12.8	4358	4	US-08-781-986A-454
15	436.8	12.5	756	4	US-09-710-279-31
16	365	10.4	396	4	US-08-956-171E-3602
17	365	10.4	396	4	US-08-781-986A-3602
18	360.6	10.3	3498	4	US-09-200-650E-8
19	360.6	10.3	4005	4	US-08-956-171E-631
20	360.6	10.3	4005	4	US-08-781-986A-631
21	359.8	10.3	2841	4	US-09-200-650E-4
22	340.8	9.7	864	4	US-09-710-279-467
23	340.8	9.7	3246	4	US-09-710-279-3336
24	340.8	9.7	3619	4	US-09-710-279-4157
25	327.6	9.4	399	4	US-08-956-171E-3555
26	327.6	9.4	399	4	US-08-781-986A-3555
27	307.6	8.8	1461	4	US-09-710-279-787

c	28	307.6	8.8	3612	4	US-09-710-279-3696	Sequence 3696, Ap
	29	303.4	8.7	400	4	US-08-956-171E-4206	Sequence 4206, Ap
	30	303.4	8.7	400	4	US-08-781-986A-4206	Sequence 4206, Ap
c	31	302.2	8.6	530	4	US-08-956-171E-3656	Sequence 3656, Ap
c	32	302.2	8.6	530	4	US-08-781-986A-3656	Sequence 3656, Ap
	33	300.8	8.6	1464	4	US-09-386-962C-13	Sequence 13, Appl
	34	300.8	8.6	1464	4	US-09-386-959-64	Sequence 64, Appl
	35	292.4	8.4	3945	4	US-09-200-650E-6	Sequence 6, Appl
c	36	292.2	8.4	573	4	US-08-956-171E-3859	Sequence 3859, Ap
c	37	292.2	8.4	573	4	US-08-781-986A-3859	Sequence 3859, Ap
c	38	272.4	7.8	476	4	US-08-956-171E-3725	Sequence 3725, Ap
c	39	272.4	7.8	476	4	US-08-781-986A-3725	Sequence 3725, Ap
	40	269.2	7.7	374	4	US-08-956-171E-4060	Sequence 4060, Ap
	41	269.2	7.7	374	4	US-08-781-986A-4060	Sequence 4060, Ap
	42	260.8	7.5	308	4	US-08-956-171E-4466	Sequence 4466, Ap
	43	260.8	7.5	308	4	US-08-781-986A-4466	Sequence 4466, Ap
c	44	250.8	7.2	400	4	US-08-956-171E-3600	Sequence 3600, Ap
c	45	250.8	7.2	400	4	US-08-781-986A-3600	Sequence 3600, Ap

ALIGNMENTS

RESULT 1
US-08-293-728-1
; Sequence 1, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293.728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3498
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (301)..(3099)
US-08-293-728-1

Query Match	100.0%;	Score 3498;	DB 3;	Length 3498;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3498;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGTACCATAAATACACATCTGCTTTTGA	AAATATGATTTTCAAGCTAGGATACATTA	60
Db	1	GGTACCATAAATACACATCTGCTTTTGA	AAATATGATTTTCAAGCTAGGATACATTA	60
QY	61	GGTAGAGTTTCATATTAATAAATAAATGTTTGC	CAATCGATCGTTTGTGTTTGTGA	120
Db	61	GGTAGAGTTTCATATTAATAAATAAATGTTTGC	CAATCGATCGTTTGTGTTTGTGA	120
QY	121	ATTCTTAAATAGCAATAAATAAATGTTTGTGTA	AGATATTTTGGATTAATAAATAA	180
Db	121	ATTCTTAAATAGCAATAAATAAATGTTTGTGTA	AGATATTTTGGATTAATAAATAA	180
QY	181	TATCGATACAAATTAATTCGATTAATTCGATTA	ATTCATTAATTCGATTAATTCGATTA	240
Db	181	TATCGATACAAATTAATTCGATTAATTCGATTA	ATTCATTAATTCGATTAATTCGATTA	240
QY	241	AAATATATCTTAAAGGGTATATAGTTTAAATAA	ATGACATTTTAAAGAGGGAATAA	300
Db	241	AAATATATCTTAAAGGGTATATAGTTTAAATAA	ATGACATTTTAAAGAGGGAATAA	300
QY	301	ATGAATATGAAGAAAAAAGAAAAACGCAATTCG	AAAAAATCGATTCGCGTTC	360
Db	301	ATGAATATGAAGAAAAAAGAAAAACGCAATTCG	AAAAAATCGATTCGCGTTC	360

QY 361 GTGCTTGAGGTACGTTAAATCGTTTTCGACTACTCAGCAGTAAAGAAAGCAGATCAAGT 420
Db 361 GTGCTTGAGGTACGTTAAATCGTTTTCGACTACTCAGCAGTAAAGAAAGCAGATCAAGT 420
QY 421 GAAATAGTGTGTTACGCAATCTGATAGCCAAAGTAAACGAAAGCAAAAGTAATGATTCAGT 480
Db 421 GAAATAGTGTGTTACGCAATCTGATAGCCAAAGTAAACGAAAGCAAAAGTAATGATTCAGT 480
QY 481 AGCGTTAGTGTGCACTTAAACAGACAGACACAAACGTTGAGTGATCTAAACATCGTCA 540
Db 481 AGCGTTAGTGTGCACTTAAACAGACAGACACAAACGTTGAGTGATCTAAACATCGTCA 540
QY 541 AACACTAATATGCGGAAACGAGTGTGCGCAAAATCCAGCAACAACGAGAAACGACAAA 600
Db 541 AACACTAATATGCGGAAACGAGTGTGCGCAAAATCCAGCAACAACGAGAAACGACAAA 600
QY 601 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGTTGAAAGTACTACTAG 660
Db 601 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGTTGAAAGTACTACTAG 660
QY 661 ACAAGCAATCAAGTAAATACACCGGCAACAACCTCAATCAAGCAATACAAATCGGAGGAA 720
Db 661 ACAAGCAATCAAGTAAATACACCGGCAACAACCTCAATCAAGCAATACAAATCGGAGGAA 720
QY 721 TTAGTGAATCAAAACAAAGTAAATGAAACGACTTTTAAATGATCTAAATACAGTATCATCTGTA 780
Db 721 TTAGTGAATCAAAACAAAGTAAATGAAACGACTTTTAAATGATCTAAATACAGTATCATCTGTA 780
QY 781 AATTCACCTCAAAATTTCTACAAATCGGAAATGTTTCAACACGCAAGATCTTCAACT 840
Db 781 AATTCACCTCAAAATTTCTACAAATCGGAAATGTTTCAACACGCAAGATCTTCAACT 840
QY 841 GAAGCAACACCTTCAAAACAAATGAAATCAAGTCCACAGAGTACAGATGCAAGTAAATAGAT 900
Db 841 GAAGCAACACCTTCAAAACAAATGAAATCAAGTCCACAGAGTACAGATGCAAGTAAATAGAT 900
QY 901 GTAGTTAATCAAGCGGTTAATCAAGTGGCCCTAGAAATGAGAGCAATTTAGTTTAGCGCA 960
Db 901 GTAGTTAATCAAGCGGTTAATCAAGTGGCCCTAGAAATGAGAGCAATTTAGTTTAGCGCA 960
QY 961 GTAGTGCAGATGCAACCGGAGTGGCAAGATATTAACGAATCAAGTTCAGCAATGTGACA 1020
Db 961 GTAGTGCAGATGCAACCGGAGTGGCAAGATATTAACGAATCAAGTTCAGCAATGTGACA 1020
QY 1021 GTTGTGATTTGACTCTGTTAGTGTGATTCGCAAGTGGTATGCTCAAACTGAT 1080
Db 1021 GTTGTGATTTGACTCTGTTAGTGTGATTCGCAAGTGGTATGCTCAAACTGAT 1080
QY 1081 TATGTTTTCAGTGCCTAATTCGTTTAAAGGTGACACATTCAAAATAAATCTGTACCT 1140
Db 1081 TATGTTTTCAGTGCCTAATTCGTTTAAAGGTGACACATTCAAAATAAATCTGTACCT 1140
QY 1141 AAGAAATTAACCTTAAATCGGTGTAATCTCAACTGCTTAAAGTGGCACCACCAATTTAGCTGGA 1200
Db 1141 AAGAAATTAACCTTAAATCGGTGTAATCTCAACTGCTTAAAGTGGCACCACCAATTTAGCTGGA 1200
QY 1201 GATCAAGTATGCGAAATCGGTGTAATCGATAGTGTGATGTTTAAATGATCTTAAATGAT 1260
Db 1201 GATCAAGTATGCGAAATCGGTGTAATCGATAGTGTGATGTTTAAATGATCTTAAATGAT 1260
QY 1261 GACTATGTAATCTAAAGATGATGTAAGAAAGCAACTTTGACCATGCGCGCTTATATTTGAC 1320
Db 1261 GACTATGTAATCTAAAGATGATGTAAGAAAGCAACTTTGACCATGCGCGCTTATATTTGAC 1320
QY 1321 CTGGAATGTTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAAACA 1380
Db 1321 CTGGAATGTTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAAACA 1380
QY 1381 GCAACAAACAGATTTAGTAGATTATGAAATATGTTAAGTATTTTAACTTATCTATT 1440
Db 1381 GCAACAAACAGATTTAGTAGATTATGAAATATGTTAAGTATTTTAACTTATCTATT 1440
QY 1441 AAGGTACAATTTGACCAAAATCGATAAAACAAATTAATACGTATCGTCAGACAAATTTATGTC 1500

Db 1441 AAGGTACAATTTGACCAAAATCGATAAAACAAATAATACGTATCGTCAGACAAATTTATGTC 1500
QY 1501 AATCCAAGTGAGATTAACGTTTATTCGCGCGGTTTTAAACAGGTAATTTAAACCAATACG 1560
Db 1501 AATCCAAGTGAGATTAACGTTTATTCGCGCGGTTTTAAACAGGTAATTTAAACCAATACG 1560
QY 1561 GATAGTAATGATTAATAGATCAGCAAAATACAAAGTATTAAGTATATAAAGTAGATAAT 1620
Db 1561 GATAGTAATGATTAATAGATCAGCAAAATACAAAGTATTAAGTATATAAAGTAGATAAT 1620
QY 1621 GCAGCTGATTTATCTGAAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGATGTCTAAT 1680
Db 1621 GCAGCTGATTTATCTGAAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGATGTCTAAT 1680
QY 1681 AGTGTGAATATTAACATTCGCAAAATCCAAATCAATATAAGTAGAGTATTAATACGCTGAT 1740
Db 1681 AGTGTGAATATTAACATTCGCAAAATCCAAATCAATATAAGTAGAGTATTAATACGCTGAT 1740
QY 1741 GATCAAAATACAAACACCGGTATATAGTAGTTGTGTTAATGCTCATATTTGATCCGAATAGCAAA 1800
Db 1741 GATCAAAATACAAACACCGGTATATAGTAGTTGTGTTAATGCTCATATTTGATCCGAATAGCAAA 1800
QY 1801 GGTGATTTAGCTTTACGTTTCAACTTTATATGGGTAACTCGATATATAATTTGGGCTCT 1860
Db 1801 GGTGATTTAGCTTTACGTTTCAACTTTATATGGGTAACTCGATATATAATTTGGGCTCT 1860
QY 1861 ATGTCATCGGCAACGAGTAGCATTTAATACCGATCAGGTTCTGTTGACGGTATCGAT 1920
Db 1861 ATGTCATCGGCAACGAGTAGCATTTAATACCGATCAGGTTCTGTTGACGGTATCGAT 1920
QY 1921 AAACAGTTGTTCCCTGAAACAACTGATGAGCTGTGAAATGAAACCAATTCAGAGGAT 1980
Db 1921 AAACAGTTGTTCCCTGAAACAACTGATGAGCTGTGAAATGAAACCAATTCAGAGGAT 1980
QY 1981 TCAGATTTGACCCAGGTTTCAGATTTGAGCGAGTTCGATTCAGATAGCGGTTTCAGAT 2040
Db 1981 TCAGATTTGACCCAGGTTTCAGATTTGAGCGAGTTCGATTCAGATAGCGGTTTCAGAT 2040
QY 2041 TCGGTTAGTGTGATTTACATCAGATAGTGGTTTCAGATTCAGCGAGTTCAGATTCAGCA 2100
Db 2041 TCGGTTAGTGTGATTTACATCAGATAGTGGTTTCAGATTCAGCGAGTTCAGATTCAGCA 2100
QY 2101 AGTGATTCAGACTCAGCGAGTTCAGATTCAGCAAGCGATTCGACTCAGCGAGCGAT 2160
Db 2101 AGTGATTCAGACTCAGCGAGTTCAGATTCAGCAAGCGATTCGACTCAGCGAGCGAT 2160
QY 2161 TCCGACTCAGCAATGATTCGGATTCAGATAGCGATTCGACTCAGCAAGTCTCAGAT 2220
Db 2161 TCCGACTCAGCAATGATTCGGATTCAGATAGCGATTCGACTCAGCAAGTCTCAGAT 2220
QY 2221 TCCGACAGTGTGATTCAGATTCAGATAGCGATTCGACTCAGCAAGTCTCAGATTCAGAT 2280
Db 2221 TCCGACAGTGTGATTCAGATTCAGATAGCGATTCGACTCAGCAAGTCTCAGATTCAGAT 2280
QY 2281 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTTCGACTCAGCAAGCGAT 2340
Db 2281 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTTCGACTCAGCAAGCGAT 2340
QY 2341 TCTGACTCCGACAGTGTTCGACTCAGCAAGCGATTCAGATTCGCAAGTTCGACTCAGCAAG 2400
Db 2341 TCTGACTCCGACAGTGTTCGACTCAGCAAGCGATTCAGATTCGCAAGTTCGACTCAGCAAG 2400
QY 2401 TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGCAAGCGATTCAGATTCAGAT 2460
Db 2401 TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGCAAGCGATTCAGATTCAGAT 2460
QY 2461 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTCTCAGATTCGCAAGTCTCAGAT 2520
Db 2461 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTCTCAGATTCGCAAGTCTCAGAT 2520
QY 2521 TCGGATTCAGATAGCGATTCAGATTCGCAAGTCTCAGATTCGCAAGTCTCAGATTCAGAT 2580

781. AATTCACTCAAAATTTCTACAAATCGGAATGTTTCAACAACGCAAGATACCTCAACT 840
Db AATTCACTCAAAATTTCTACAAATCGGAATGTTTCAACAACGCAAGATACCTCAACT 840
781. AATTCACTCAAAATTTCTACAAATCGGAATGTTTCAACAACGCAAGATACCTCAACT 840
781. AATTCACTCAAAATTTCTACAAATCGGAATGTTTCAACAACGCAAGATACCTCAACT 840
841. GAAGCAACACCTTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT 900
Db GAAGCAACACCTTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT 900
901. GTAGTTAATCAAGCGGTTAATACAGTGGGCTAGAAATGAGAGCAATTTAGTTAGCGCA 960
Db GTAGTTAATCAAGCGGTTAATACAGTGGGCTAGAAATGAGAGCAATTTAGTTAGCGCA 960
901. GTAGTTAATCAAGCGGTTAATACAGTGGGCTAGAAATGAGAGCAATTTAGTTAGCGCA 960
901. GTAGTTAATCAAGCGGTTAATACAGTGGGCTAGAAATGAGAGCAATTTAGTTAGCGCA 960
961. GTAGCTGAGATGCAACCGGAGCTGGCAAGATATTTACGAATCAGTTGACGAAATGTGACA 1020
Db GTAGCTGAGATGCAACCGGAGCTGGCAAGATATTTACGAATCAGTTGACGAAATGTGACA 1020
961. GTAGCTGAGATGCAACCGGAGCTGGCAAGATATTTACGAATCAGTTGACGAAATGTGACA 1020
961. GTAGCTGAGATGCAACCGGAGCTGGCAAGATATTTACGAATCAGTTGACGAAATGTGACA 1020
1021. GTTGTATTTGACTCTGGTACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAAATGAAT 1080
Db GTTGTATTTGACTCTGGTACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAAATGAAT 1080
1021. GTTGTATTTGACTCTGGTACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAAATGAAT 1080
1021. GTTGTATTTGACTCTGGTACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAAATGAAT 1080
1081. TATGGTTTTTTCAGTGCCTAAATTCCTGCTGTTAAAGGTGACATTTCAAAATTAATCTGACCT 1140
Db TATGGTTTTTTCAGTGCCTAAATTCCTGCTGTTAAAGGTGACATTTCAAAATTAATCTGACCT 1140
1081. TATGGTTTTTTCAGTGCCTAAATTCCTGCTGTTAAAGGTGACATTTCAAAATTAATCTGACCT 1140
1081. TATGGTTTTTTCAGTGCCTAAATTCCTGCTGTTAAAGGTGACATTTCAAAATTAATCTGACCT 1140
1141. AAAGAAATTAACCTTAAATGGTGAATTTCAACTGCTTAAAGTCCACCAATTTATGGCTGA 1200
Db AAAGAAATTAACCTTAAATGGTGAATTTCAACTGCTTAAAGTCCACCAATTTATGGCTGA 1200
1141. AAAGAAATTAACCTTAAATGGTGAATTTCAACTGCTTAAAGTCCACCAATTTATGGCTGA 1200
1141. AAAGAAATTAACCTTAAATGGTGAATTTCAACTGCTTAAAGTCCACCAATTTATGGCTGA 1200
1201. GATCAAGTATTTGGCAAAATGGTGAATTCGATAGTGAATGTTATTTATTTATTTATTTACA 1260
Db GATCAAGTATTTGGCAAAATGGTGAATTCGATAGTGAATGTTATTTATTTATTTATTTACA 1260
1201. GATCAAGTATTTGGCAAAATGGTGAATTCGATAGTGAATGTTATTTATTTATTTATTTACA 1260
1201. GATCAAGTATTTGGCAAAATGGTGAATTCGATAGTGAATGTTATTTATTTATTTATTTACA 1260
1261. GACTATGTAATTAAGATGATTAAGAAAGCAATTTGACCATCGCCGCTTATATTGAC 1320
Db GACTATGTAATTAAGATGATTAAGAAAGCAATTTGACCATCGCCGCTTATATTGAC 1320
1261. GACTATGTAATTAAGATGATTAAGAAAGCAATTTGACCATCGCCGCTTATATTGAC 1320
1261. GACTATGTAATTAAGATGATTAAGAAAGCAATTTGACCATCGCCGCTTATATTGAC 1320
1321. CCTGAAATTTGTTAAAGACAGTGAATGTGACATTTGGCTACTGCGCATAGGTAGTACAACA 1380
Db CCTGAAATTTGTTAAAGACAGTGAATGTGACATTTGGCTACTGCGCATAGGTAGTACAACA 1380
1321. CCTGAAATTTGTTAAAGACAGTGAATGTGACATTTGGCTACTGCGCATAGGTAGTACAACA 1380
1321. CCTGAAATTTGTTAAAGACAGTGAATGTGACATTTGGCTACTGCGCATAGGTAGTACAACA 1380
1381. GCAAAACAAACAGTATTTAGTATGATTTGAAAATATGGTAACTTTTAACTTTATCTATT 1440
Db GCAAAACAAACAGTATTTAGTATGATTTGAAAATATGGTAACTTTTAACTTTATCTATT 1440
1381. GCAAAACAAACAGTATTTAGTATGATTTGAAAATATGGTAACTTTTAACTTTATCTATT 1440
1381. GCAAAACAAACAGTATTTAGTATGATTTGAAAATATGGTAACTTTTAACTTTATCTATT 1440
1441. AAAGTACAATTTGACCAATCGATTAAGAAACAAATTAATGATATCGTCAAGCAATTTATGTC 1500
Db AAAGTACAATTTGACCAATCGATTAAGAAACAAATTAATGATATCGTCAAGCAATTTATGTC 1500
1441. AAAGTACAATTTGACCAATCGATTAAGAAACAAATTAATGATATCGTCAAGCAATTTATGTC 1500
1441. AAAGTACAATTTGACCAATCGATTAAGAAACAAATTAATGATATCGTCAAGCAATTTATGTC 1500
1501. AATCCAAAGTGAATTAACCTTTATTCGCGCGGTTTAAACAGGTAATTTTAAACCAAAATACG 1560
Db AATCCAAAGTGAATTAACCTTTATTCGCGCGGTTTAAACAGGTAATTTTAAACCAAAATACG 1560
1501. AATCCAAAGTGAATTAACCTTTATTCGCGCGGTTTAAACAGGTAATTTTAAACCAAAATACG 1560
1501. AATCCAAAGTGAATTAACCTTTATTCGCGCGGTTTAAACAGGTAATTTTAAACCAAAATACG 1560
1561. GATAGTAATGCAATTAATAGATCAGCAAAATACAGATTTTAAAGTATATAAGTAGATAAT 1620
Db GATAGTAATGCAATTAATAGATCAGCAAAATACAGATTTTAAAGTATATAAGTAGATAAT 1620
1561. GATAGTAATGCAATTAATAGATCAGCAAAATACAGATTTTAAAGTATATAAGTAGATAAT 1620
1561. GATAGTAATGCAATTAATAGATCAGCAAAATACAGATTTTAAAGTATATAAGTAGATAAT 1620
1621. GCAGCTGATTTATCTGAAAGTTACTTTGTGATCCAGAAACTTTGAGGATGTCATAAT 1680
Db GCAGCTGATTTATCTGAAAGTTACTTTGTGATCCAGAAACTTTGAGGATGTCATAAT 1680
1621. GCAGCTGATTTATCTGAAAGTTACTTTGTGATCCAGAAACTTTGAGGATGTCATAAT 1680
1621. GCAGCTGATTTATCTGAAAGTTACTTTGTGATCCAGAAACTTTGAGGATGTCATAAT 1680
1681. AGTGTGAATATTACATTTCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCC 1740
Db AGTGTGAATATTACATTTCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCC 1740
1681. AGTGTGAATATTACATTTCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCC 1740
1681. AGTGTGAATATTACATTTCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCC 1740
1741. GATCAAAATTAACACCGGATATAGTATGTTTAAATGTTATGTTATGTTATGTTATGTTATGTT 1800
Db GATCAAAATTAACACCGGATATAGTATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTT 1800
1741. GATCAAAATTAACACCGGATATAGTATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTT 1800
1741. GATCAAAATTAACACCGGATATAGTATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTT 1800
1801. GGTGATTTAGCTTTTACGTTCACTTTATGTTGTTATGTTTAAATGTTTAAATGTTTAAATGTTT 1860
Db GGTGATTTAGCTTTTACGTTCACTTTATGTTGTTATGTTTAAATGTTTAAATGTTTAAATGTTT 1860
1801. GGTGATTTAGCTTTTACGTTCACTTTATGTTGTTATGTTTAAATGTTTAAATGTTTAAATGTTT 1860
1801. GGTGATTTAGCTTTTACGTTCACTTTATGTTGTTATGTTTAAATGTTTAAATGTTTAAATGTTT 1860

1861. ATGTCATGGGCAACGAAGTAGCTATTTAATAACCGATCAGGTTCTGGTACCGTATCGAT 1920
Db ATGTCATGGGCAACGAAGTAGCTATTTAATAACCGATCAGGTTCTGGTACCGTATCGAT 1920
1861. ATGTCATGGGCAACGAAGTAGCTATTTAATAACCGATCAGGTTCTGGTACCGTATCGAT 1920
1861. ATGTCATGGGCAACGAAGTAGCTATTTAATAACCGATCAGGTTCTGGTACCGTATCGAT 1920
1921. AAACAGTGTGTTCTGTGAACAACTGTAGTGGCTGTGTGAAAATTTGAACCAATTTCCAGAGAT 1980
Db AAACAGTGTGTTCTGTGAACAACTGTAGTGGCTGTGTGAAAATTTGAACCAATTTCCAGAGAT 1980
1921. AAACAGTGTGTTCTGTGAACAACTGTAGTGGCTGTGTGAAAATTTGAACCAATTTCCAGAGAT 1980
1921. AAACAGTGTGTTCTGTGAACAACTGTAGTGGCTGTGTGAAAATTTGAACCAATTTCCAGAGAT 1980
1981. TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGACGAGTTCTAATTCAGATAGCGGTTTCAGAT 2040
Db TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGACGAGTTCTAATTCAGATAGCGGTTTCAGAT 2040
1981. TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGACGAGTTCTAATTCAGATAGCGGTTTCAGAT 2040
1981. TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGACGAGTTCTAATTCAGATAGCGGTTTCAGAT 2040
2041. TCGGCTAGTGTATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA 2100
Db TCGGCTAGTGTATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA 2100
2041. TCGGCTAGTGTATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA 2100
2041. TCGGCTAGTGTATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA 2100
2101. AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGAT 2160
Db AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGAT 2160
2101. AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGAT 2160
2101. AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGAT 2160
2161. TCCGACTCAGACAATGACTTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGAT 2220
Db TCCGACTCAGACAATGACTTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGAT 2220
2161. TCCGACTCAGACAATGACTTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGAT 2220
2161. TCCGACTCAGACAATGACTTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGAT 2220
2221. TCCGACAGTGAATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGATTCAGAT 2280
Db TCCGACAGTGAATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGATTCAGAT 2280
2221. TCCGACAGTGAATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGATTCAGAT 2280
2221. TCCGACAGTGAATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGATTCAGAT 2280
2281. AGCGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGAT 2340
Db AGCGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGAT 2340
2281. AGCGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGAT 2340
2281. AGCGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGAT 2340
2341. TCTGACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGACAGTGAATTCGAC 2400
Db TCTGACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGACAGTGAATTCGAC 2400
2341. TCTGACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGACAGTGAATTCGAC 2400
2341. TCTGACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGACAGTGAATTCGAC 2400
2401. TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGAT 2460
Db TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGAT 2460
2401. TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGAT 2460
2401. TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGAT 2460
2461. AGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACAGTGAATTCAGATTCGCGACAGTGAATTCAGAT 2520
Db AGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACAGTGAATTCAGATTCGCGACAGTGAATTCAGAT 2520
2461. AGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACAGTGAATTCAGATTCGCGACAGTGAATTCAGAT 2520
2461. AGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACAGTGAATTCAGATTCGCGACAGTGAATTCAGAT 2520
2521. TCGGATTCAGATAGCGATTCAGATTCGCGACAGTGAATTCAGATTCGCGACAGTGAATTCAGATTCAGAT 2580
Db TCGGATTCAGATAGCGATTCAGATTCGCGACAGTGAATTCAGATTCGCGACAGTGAATTCAGATTCAGAT 2580
2521. TCGGATTCAGATAGCGATTCAGATTCGCGACAGTGAATTCAGATTCGCGACAGTGAATTCAGATTCAGAT 2580
2521. TCGGATTCAGATAGCGATTCAGATTCGCGACAGTGAATTCAGATTCGCGACAGTGAATTCAGATTCAGAT 2580
2581. TCAGACAGTGAATTCGCGATTCAGCGAGTGAATTCGCGATTCAGATTCGCGATTCGCGATTCGCGATTCGCGAT 2640
Db TCAGACAGTGAATTCGCGATTCAGCGAGTGAATTCGCGATTCAGATTCGCGATTCGCGATTCGCGATTCGCGAT 2640
2581. TCAGACAGTGAATTCGCGATTCAGCGAGTGAATTCGCGATTCAGATTCGCGATTCGCGATTCGCGATTCGCGAT 2640
2581. TCAGACAGTGAATTCGCGATTCAGCGAGTGAATTCGCGATTCAGATTCGCGATTCGCGATTCGCGATTCGCGAT 2640
2641. AGTGATTCGCGATTCAGATAGCGATTCAGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGAT 2700
Db AGTGATTCGCGATTCAGATAGCGATTCAGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGAT 2700
2641. AGTGATTCGCGATTCAGATAGCGATTCAGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGAT 2700
2641. AGTGATTCGCGATTCAGATAGCGATTCAGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGAT 2700
2701. TCGGATTCAGATAGCGATTCAGATTCGCGATTCAGATTCGCGATTCAGATTCGCGATTCGCGATTCGCGAT 2760
Db TCGGATTCAGATAGCGATTCAGATTCGCGATTCAGATTCGCGATTCAGATTCGCGATTCGCGATTCGCGAT 2760
2701. TCGGATTCAGATAGCGATTCAGATTCGCGATTCAGATTCGCGATTCAGATTCGCGATTCGCGATTCGCGAT 2760
2701. TCGGATTCAGATAGCGATTCAGATTCGCGATTCAGATTCGCGATTCAGATTCGCGATTCGCGATTCGCGAT 2760
2761. TCAGACAGCGATTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2820
Db TCAGACAGCGATTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2820
2761. TCAGACAGCGATTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2820
2761. TCAGACAGCGATTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2820
2821. TCAGACTCAGGTAGTGAATTCGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2880
Db TCAGACTCAGGTAGTGAATTCGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2880
2821. TCAGACTCAGGTAGTGAATTCGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2880
2821. TCAGACTCAGGTAGTGAATTCGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2880
2881. AGCGATTCGCGATTCAGGTTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGAT 2940
Db AGCGATTCGCGATTCAGGTTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGAT 2940
2881. AGCGATTCGCGATTCAGGTTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGAT 2940
2881. AGCGATTCGCGATTCAGGTTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGATTCGCGAT 2940
2941. ACTAATGCTTCTAATAAANAAGAGCTAAGATAGTAAAGACCATTTACCATACAGAT 3000
Db ACTAATGCTTCTAATAAANAAGAGCTAAGATAGTAAAGACCATTTACCATACAGAT 3000

Db 2941 ACTAATGCTTCTTAATAAATAGGCTAAAGATAGTAAGAACCACTTACCAGATACAGGT 3000
QY 3001 TCTGAGATGAAGCAATAGCTACTAATTTGGGATATTAGCATCAATAGCTTCAITA 3060
Db 3001 TCTGAGATGAAGCAATAGCTACTAATTTGGGATATTAGCATCAATAGCTTCAITA 3060
QY 3061 CTACTTTTCAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 3120
Db 3061 CTACTTTTCAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 3120
QY 3121 TAATCATATGATTCATGAAGAGCCACCTTAAAGAGGTCCTTTTACTTGGATTTTCCA 3180
Db 3121 TAATCATATGATTCATGAAGAGCCACCTTAAAGAGGTCCTTTTACTTGGATTTTCCA 3180
QY 3181 AATATATGTTTGAATATTAATTAATTAATTCATCAACAGTTAAATTTTAAAAAGG 3240
Db 3181 AATATATGTTTGAATATTAATTAATTCATCAACAGTTAAATTTTAAAAAGG 3240
QY 3241 TAGATGTTATATATTTGGCTTGGCGAAAAAATAGGCTGTAAGGTAGGTGTTTAAATTAGG 3300
Db 3241 TAGATGTTATATTTGGCTTGGCGAAAAAATAGGCTGTAAGGTAGGTGTTTAAATTAGG 3300
QY 3301 GAAAAATTAAGAGAAAAAATACAGTTGAAAAAATAAATTCCTAGTTTATCATTTGGGAGCATT 3360
Db 3301 GAAAAATTAAGAGAAAAAATACAGTTGAAAAAATAAATTCCTAGTTTATCATTTGGGAGCATT 3360
QY 3361 ATGCTATCAAAATTTGGGAAAGTAATCGTGGAGTGCAGTGGTTTCTGGGAGAGAA 3420
Db 3361 ATGCTATCAAAATTTGGGAAAGTAATCGTGGAGTGCAGTGGTTTCTGGGAGAGAA 3420
QY 3421 TCCATATGATCTGAGTCGTTGAACTGACTATAATAAAAAAATAAACTTAGAACAGTAGA 3480
Db 3421 TCCATATGATCTGAGTCGTTGAACTGACTATAATAAAAAAATAAACTTAGAACAGTAGA 3480
QY 3481 AGAGTATAAGAAAGCTT 3498
Db 3481 AGAGTATAAGAAAGCTT 3498

RESULT 3

US-08-956-171E-587
; Sequence 587, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956.171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PE248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 587:
SEQUENCE CHARACTERISTICS:
LENGTH: 4709 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 587:
US-08-956-171E-587

Query Match 97.2%; Score 3399.6; DB 4; Length 4709;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3466; Conservative 1; Mismatches 11; Indels 21; Gaps 4;
QY 1 GGTACCAATAAATACACATCTGCTTTTGAATAAATATGATTTCAGCTAGATTACATT 60
Db 1204 GGTACCAATAAATACACATCTGCTTTTGAATAAATATGATTTCAGCTAGATTACATT 1263
QY 61 GGTAGAGTTCATATTAATAAATAAATGTTTGCATCAATCGTACGTTGTCGTTGTA 120
Db 1264 GGTAGAGTTCATATTAATAAATAAATGTTTGCATCAATCGTACGTTGTCGTTGTA 1323
QY 121 ATTCCTTAAAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTGTGATAATAAAA 180
Db 1324 ATTCCTTAAAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTGTGATAATAAAA 1383
QY 181 TATCGATACAAATTAATGCTATTAATGCAATTTTGTAGTATTAATTCATTAAACAGAGATT 240
Db 1384 TATCGATACAAATTAATGCTATTAATGCAATTTTGTAGTATTAATTCATTAAACAGAGATT 1443
QY 241 AATATATATC-TTAAAGGGTATATAGTTAATATAAATAAGCTTTTAAATAAGAGGGATAA 299
Db 1444 AATATATATC-TTAAAGGGTATATAGTTAATATAAATAAGCTTTTAAATAAGAGGGATAA 1503
QY 300 AATGAATATGAAGAAAAAAGAAAAAAGCAATTCGGAATAAATTCGATTCGGCTGGCTTC 359
Db 1504 AATGAATATGAAGAAAAAAGAAAAAAGCAATTCGGAATAAATTCGATTCGGCTGGCTTC 1563
QY 360 AGTCTTTGATGAGTACGTTAATCGGTTTGGACATCTCAGCAGTAAGAAGACAGATGCAAG 419
Db 1564 AGTCTTTGATGAGTACGTTAATCGGTTTGGACATCTCAGCAGTAAGAAGACAGATGCAAG 1623
QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAGCAAAAGTAAATGATTCAAG 479
Db 1624 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAGCAAAAGTAAATGATTCAAG 1683
QY 480 TAGCGTTAGTCTGCACCTTAAACAGACACAAACGTCGAGTGATCTACTAAAAACATCGTC 539
Db 1684 TAGCGTTAGTCTGCACCTTAAACAGACACAAACGTCGAGTGATCTACTAAAAACATCGTC 1743
QY 540 AAACACTTAATAATGGCGAAACGAGTGTGGCGAAAAATCCAGCAACACAGGAAACGACACA 599
Db 1744 AAACACTTAATAATGGCGAAACGAGTGTGGCGAAAAATCCAGCAACACAGGAAACGACACA 1803
QY 600 ATCATCATCAACAAATGCACTACGGAAGAAACGCGGTAACTGTTGCAAGCTACTACTAC 659
Db 1804 ATCATCATCAACAAATGCACTACGGAAGAAACGCGGTAACTGTTGCAAGCTACTACTAC 1863
QY 660 GACAAAGCAATCAAGCTAATACACCGCAACAACTCAATCAAGCAATCAAAATCGGAGGA 719
Db 1864 GACAAAGCAATCAAGCTAATACACCGCAACAACTCAATCAAGCAATCAAAATCGGAGGA 1923
QY 720 ATTAGTGAATCAAAACAGTAATGAAACGACTTTTAAATGATCTAATAACAGTATCATCTGT 779
Db 1924 ATTAGTGAATCAAAACAGTAATGAAACGACTTTTAAATGATCTAATAACAGTATCATCTGT 1983

780 QY AAATTCACTCAAAATTTCAAAATGCGGAAATGTTTCAACACGCAAGATACCTTCAAC 839
1984 Db AAATTCACTCAAAATTTCTAAATGCGGAAATGTTTCAACACGCAAGATACCTTCAAC 2043
840 QY TGAACCAACACCTTCAAAATGAATAGCTCCACAGAGTACAGATGCAAGTAAATGAAGA 899
2044 Db TGAACCAACACCTTCAAAATGAATAGCTCCACAGAGTACAGATGCAAGTAAATGAAGA 2103
900 QY TGTAGTTAATCAAGCGGTTAATACAAAGTCCGCTAGAAATGAGATGAGATTTAGTTCAGCGC 959
2104 Db TGTAGTTAATCAAGCGGTTAATACAAAGTCCGCTAGAAATGAGATGAGATTTAGTTCAGCGC 2163
960 QY AGTAGCTGACAGTACACCGGAGCTGGACAGATATTACGAATCAGTTGACGAAATGTGAC 1019
2164 Db AGTAGCTGACAGTACACCGGAGCTGGACAGATATTACGAATCAGTTGACGAAATGTGAC 2223
1020 QY AGTTGGTATTTGACTCTGGTACGACTGTGTATCCGACCAAGCAGGTTATGTCAAACTGAA 1079
2224 Db AGTTGGTATTTGACTCTGGTACGACTGTGTATCCGACCAAGCAGGTTATGTCAAACTGAA 2283
1080 QY TTATGGTATTTTCAAGTCCCTTAATCTGCTGTTAAAGGTGACACATTCAAAATAACTGTACC 1139
2284 Db TTATGGTATTTTCAAGTCCCTTAATCTGCTGTTAAAGGTGACACATTCAAAATAACTGTACC 2343
1140 QY TAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTGCACCAATTTATGGCTGG 1199
2344 Db TAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTGCACCAATTTATGGCTGG 2403
1200 QY AGATCAAGTATTTGGCAATGGTGTAAATCGATAGTGAATGTTATTTATACATTTAC 1259
2404 Db AGATCAAGTATTTGGCAATGGTGTAAATCGATAGTGAATGTTATTTATACATTTAC 2463
1260 QY AGACTATGTAAATACATAAGATGATTAAGCACTTTGACCATGCGCGCTTATATTGA 1319
2464 Db AGACTATGTAAATACATAAGATGATTAAGCACTTTGACCATGCGCGCTTATATTGA 2523
1320 QY CCTCGAAATGTTAAAGACAGAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAAC 1379
2524 Db CCTCGAAATGTTAAAGACAGAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAAC 2583
1380 QY AGCAAAACAAAACAGTATTTAGTATGATTAAGAAAATATGTTAAGTTTATTAATCTTAT 1439
2584 Db AGCAAAACAAAACAGTATTTAGTATGATTAAGAAAATATGTTAAGTTTATTAATCTTAT 2643
1440 QY TAAAGGTACAAATGACCAATCGATAAACAATAATAGTATCGTTCAGACAAATTTATGT 1499
2644 Db TAAAGGTACAAATGACCAATCGATAAACAATAATAGTATCGTTCAGACAAATTTATGT 2703
1500 QY CAATCCAGTGGAGATAACGTTTATTCGCGCGGTTTAAACAGGTAAATTTAAAAACCAATAC 1559
2704 Db CAATCCAGTGGAGATAACGTTTATTCGCGCGGTTTAAACAGGTAAATTTAAAAACCAATAC 2763
1560 QY GGATAGTAAATGATTAATAGATCAGCAAAATCAAGTATTAAGTATATATAAGTAGATAA 1619
2764 Db GGATAGTAAATGATTAATAGATCAGCAAAATCAAGTATTAAGTATATATAAGTAGATAA 2823
1620 QY TGCAGCTGATTTATCTGAAAGTTACTTGTGAAATCCAGAAAACCTTTGAGGATGTCACATA 1679
2824 Db TGCAGCTGATTTATCTGAAAGTTACTTGTGAAATCCAGAAAACCTTTGAGGATGTCACATA 2883
1680 QY TAGTGTGAATTTATACATTTCCCAATCCAAATCAATATATAAGTAGAGTTTAAATACGCTGA 1739
2884 Db TAGTGTGAATTTATACATTTCCCAATCCAAATCAATATATAAGTAGAGTTTAAATACGCTGA 2943
1740 QY TGATCAAAATTAACACCGGTATATAGTATGTTTAAATGTCATATTTGATTCGAAATAGCAA 1799
2944 Db TGATCAAAATTAACACCGGTATATAGTATGTTTAAATGTCATATTTGATTCGAAATAGCAA 3003
1800 QY AGGTGATTTTACGTTTACGTTTATATGGGTATTAATCGAATATATTTTGGCGCTC 1859
3004 Db AGGTGATTTTACGTTTATATGGGTATTAATCGAATATATTTTGGCGCTC 3063
1860 QY TATGTCATGGGACACGAAGTAGCATTTAATAACGGATCAGGTTCTGTGACGGTATCGA 1919

3064 Db TATGTCATGGGACACGAAGTAGCATTTAATAACGGATCAGGTTCTGTGTCAGGTTATCGA 3123
1920 QY TAAACCAAGTTGTTCTCTGAAACCACTGATGAGCTGTTGAAATTTGAACCAATTTCCAGAGA 1979
3124 Db TAAACCAAGTTGTTCTCTGAAACCACTGATGAGCTGTTGAAATTTGAACCAATTTCCAGAGA 3183
1980 QY TTCAGATTCTCACCCAGGTTTCAGATTCTGGCAGCGGATTTCTAATTCAGATAGCGGTTTCA 2039
3184 Db TTCAGATTCTCACCCAGGTTTCAGATTCTGGCAGCGGATTTCTAATTCAGATAGCGGTTTCA 3243
2040 QY TTCGGGTAGTGAATTTACATCAGATAGTGGTTTCAGATTTCAGCGAGTGAATTCAGATTCAGC 2099
3244 Db TTCGGGTAGTGAATTTCTCATCAGATAGTGGTTTCAGATTTCAGCGAGTGAATTCAGATTCAGC 3303
2100 QY AAGTGATTTACAGCTCAGCGAGTGAATTCAGATTTCAGCAAGCGGATTTCCGACTCAGCGGAGA 2159
3304 Db AAGTGATTTACAGCTCAGCGAGTGAATTCAGATTTCAGCAAGCGGATTTCCGACTCAGCGGAGA 3363
2160 QY TTTCCGACTCAGACAAATGACTCGGATTCAGATAGCGGATTTCTGACTCAGACAGTGAATTCAGA 2219
3364 Db TTTCCGACTCAGACAAATGACTCGGATTCAGATAGCGGATTTCTGACTCAGACAGTGAATTCAGA 3423
2220 QY TTTCCGACAGTGAATTCAGATTTCAGATAGCGGATTTCTGACTCAGACAGTGAATTCAGA 2279
3424 Db TTTCCGACAGTGAATTCAGATTTCAGATAGCGGATTTCTGACTCAGACAGTGAATTCAGA 3483
2280 QY TAGCGATTTACAGATTTCAGATAGCGGATTTCCGACAGTGAATTTCCGACTCAGACAGGAGA 2339
3484 Db TAGCGATTTACAGATTTCAGATAGCGGATTTCCGACAGTGAATTTCCGACTCAGACAGGAGA 3543
2340 QY TTTCTGACTCCGACAGTGAATTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTGAATTCAGA 2399
3544 Db TTTCTGACTCCGACAGTGAATTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTGAATTCAGA 3603
2400 QY CTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGACTCAGACAGCGGATTTAGATTCAGA 2459
3604 Db CTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGACTCAGACAGCGGATTTAGATTCAGA 3663
2460 QY CAGCGATTTACAGATTTCAGATAGCGGATTTCCGACTCAGACAGTGAATTTCCGACTCAGACAGTGA 2519
3664 Db CAGCGATTTACAGATTTCAGATAGCGGATTTCCGACTCAGACAGTGAATTTCCGACTCAGACAGTGA 3723
2520 QY CTCGAGATTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGACTCAGACAGTGAATTCAGA 2579
3724 Db CTCGAGATTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGACTCAGACAGTGAATTCAGA 3783
2580 QY CTCAGACAGTGAATTCGGAATTCAGCGAGTGAATTCGGAATTCAGATAGTGAATTCGCACTCCGA 2639
3784 Db CTCAGACAGTGAATTCGGAATTCAGCGAGTGAATTCGGAATTCAGATAGTGAATTCGCACTCCGA 3843
2640 QY CAGTGACTTCGGAATTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGACTCAGATAGCGA 2699
3844 Db CAGTGACTTCGGAATTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGACTCAGATAGCGA 3903
2700 QY TTTCCGACTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGA 2759
3904 Db TTTCCGACTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGACTCAGATAGCGGATTTCCGA 3945
2760 QY TTTCCGACAGCGGATTCAGACAGTGAATTCAGATAGTGAATTCGGAATTCAGCGGATGA 2819
3946 Db TTTCCGACAGCGGATTCAGACAGTGAATTCAGATAGTGAATTCGGAATTCAGCGGATGA 4005
2820 QY TTTCCGACTCAGTGAATTCGGAATTCAGTGAATTTCCGACTCAGAAAGTGAATTTCAAA 2879
4006 Db TTTCCGACTCAGTGAATTCGGAATTCAGTGAATTTCCGACTCAGAAAGTGAATTTCAAA 4065
2880 QY TAGCGATTTCCGAGTTCAGGTTCTTAAACAAATAGTGAATTTCCGCTAATTTCACTTAAATAGG 2939
4066 Db TAGCGATTTCCGAGTTCAGGTTCTTAAACAAATAGTGAATTTCCGCTAATTTCACTTAAATAGG 4125
2940 QY TACTAATGCTTCTAATAAATAAGGCTAAAGATAGTAAAGACCAATTTACCGATACAGG 2999

Db 4126 TACTAATGCTTCTTAATAAAAAATGAGGCTAAAGATAGTAAAGAACCACTTACCAGATACAGG 4185
QY 3000 TTCTGAGATGAGCAAAATACGTCACATAATTTGGGATTAATACCATCAATAGTTCAAT 3059
Db 4186 TTCTGAGATGAGCAAAATACGTCACATAATTTGGGATTAATACCATCAATAGTTCAAT 4245
QY 3060 ACTACTTTTTCAGAGAGAAAAAGAAAAATAAGATAGTAAAGATAGTAAATATATATAA 3119
Db 4246 ACTACTTTTTCAGAGAGAAAAAGAAAAATAAGATAGTAAAGATAGTAAATATATAA 4305
QY 3120 TTAATCATATGATTCATGAGAGAGCCACCTTAAAGGTGCTCTTTTACCTTGATTTTCC 3179
Db 4306 TTAATCATATGATTCATGAGAGAGNA-ACRCCTTAAAGGTGCTCTTTTACCTTGATTTTCC 4364
QY 3180 AATATATATGTTTGAATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3239
Db 4365 AATATATATGTTTGAATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4424
QY 3240 GTAGATGTTTATATAATTTGGCTTTGGCGAAAAATAGGGTGTAAAGGTAGGTTGTTAATTAG 3299
Db 4425 GTAGATGTTTATATAATTTGGCTTTGGCGAAAAATAGGGTGTAAAGGTAGGTTGTTAATTAG 4484
QY 3300 GGAATAATTAAGAGAGAAAAATACAGTTGAAAAATAAATTTGCTAGTTTATCATTTGGGAGCAT 3359
Db 4485 GGAATAATTAAGAGAGAAAAATACAGTTGAAAAATAAATTTGCTAGTTTATCATTTGGGAGCAT 4544
QY 3360 TATGTGTATCACAAATTTGGGAAAGTAAATCGTGCAGTCAGTGGTCTTCTGGGAGAGAA 3419
Db 4545 TATGTGTATCACAAATTTGGGAAAGTAAATCGTGCAGTCAGTGGTCTTCTGGGAGAGAA 4604
QY 3420 ATCCATATGATCTGAGTCTGTTGAAACTGACTAATAATAAAAAATFAATCTAGAACAGTAG 3479
Db 4605 ATCCATATGATCT-AGTCTGTTGAACTGACTAATAATAAAAAATCTAGAACAGTAG 4663
QY 3480 AGAGATATAGAAAAGCTT 3498
Db 4664 AAGAGTATAGAAAAGATT 4682

RESULT 4

US-08-781-986A-587
; Sequence 587, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 587:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-587

Query Match 97.2%; Score 3399.6; DB 4; Length 4709;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3466; Conservative 1; Mismatches 11; Indels 21; Gaps 4;

QY 1 GGTACCATAAATACACATCTGCTTTTGAATAAATATGATTTCAAGTAGGATTCATATTA 60
Db 1204 GGTACCATAAATACACATCTGCTTTTGAATAAATATGATTTCAAGTAGGATTCATATTA 1263
QY 61 GGTAGAGTTTCATATTAATAATAAAAAATGTTTGCATCAATCGTAGCTTGTCTTTGTA 120
Db 1264 GGTAGAGTTTCATATTAATAATAAAAAATGTTTGCATCAATCGTAGCTTGTCTTTGTA 1323
QY 121 ATTCTTAAATAGCAATAATAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA 180
Db 1324 ATTCTTAAATAGCAATAATAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA 1383
QY 181 TATCGATACAAAATTAATTTGCTATAATGCAATTTTAGTGTATAATTCATTTAAACAGAGATT 240
Db 1384 TATCGATACAAAATTAATTTGCTATAATGCAATTTTAGTGTATAATTCATTTAAACAGAGATT 1443
QY 241 AATATATATC-TTAAAGGTATATAGTTTAAATATAATAATGACTTTTTTAAAAAGAGGGAATAA 299
Db 1444 AATATATATCTTTAAAGGTATATAGTTTAAATATAATAATGACTTTTTTAAAAAGAGGGAATAA 1503
QY 300 AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGCGCTGGCTTC 359
Db 1504 AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGCGCTGGCTTC 1563
QY 360 AGTGCTTTAGGTACGTTTAAATCGGTTTGGACTACTCAGCAGTAAAGAACAGATGCAAG 419
Db 1564 AGTGCTTTAGGTACGTTTAAATCGGTTTGGACTACTCAGCAGTAAAGAACAGATGCAAG 1623
QY 420 TGAATAATAGTTTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAAAGTAAATGATTCAAG 479
Db 1624 TGAATAATAGTTTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAAAGTAAATGATTCAAG 1683
QY 480 TAGCGTTAGTGTGTCACCTTAAACACAGCAGCACAACGCTGAGTGATCTAAACATCGTC 539
Db 1684 TAGCGTTAGTGTGTCACCTTAAACACAGCAGCACAACGCTGAGTGATCTAAACATCGTC 1743
QY 540 AAACACTTAATAATGGCGAAACAGAGTGTGGCGCAAAATCCAGCACAACAGGAAACGACACA 599
Db 1744 AAACACTTAATAATGGCGAAACAGAGTGTGGCGCAAAATCCAGCACAACAGGAAACGACACA 1803
QY 600 ATCATCATCAACAAATGCAACTACGGAAGAAACCGCGTAACTGGTGAAGCTACTACTAC 659
Db 1804 ATCATCATCAACAAATGCAACTACGGAAGAAACCGCGTAACTGGTGAAGCTACTACTAC 1863
QY 660 GACAAAGAAATCAGCTTAATACACCGGCAACACTCAATCAGCAATACAAATCGGAGGA 719
Db 1864 GACAAAGAAATCAGCTTAATACACCGGCAACACTCAATCAGCAATACAAATCGGAGGA 1923
QY 720 ATTAGTGAATCAAAACAAAGTAAATGAAACGACTTTTAATGATCTAATACAGTATCATCTGT 779
Db 1924 ATTAGTGAATCAAAACAAAGTAAATGAAACGACTTTTAATGATCTAATACAGTATCATCTGT 1983
QY 780 AATTCACCTTCAAAATTTCAAAATCGGAAAAATGTTTCAACAAACGCAAGATACTTCAAC 839
Db 1984 AATTCACCTTCAAAATTTCAAAATCGGAAAAATGTTTCAACAAACGCAAGATACTTCAAC 2043
QY 840 TGAAGCAACACCTTCAACAAATGAAATCAGCTCCACAGAGTACAGATGCAAGTAAATAAGA 899
Db 2044 TGAAGCAACACCTTCAACAAATGAAATCAGCTCCACAGAGTACAGATGCAAGTAAATAAGA 2103
QY 900 TGTAGTTAATCAAGCGGTTTAAATACAAAGTGGCGCTAGATGAGAGCATTTTAGTTTACGGCG 959

Best Local Similarity 55.2%; Pred. No. 2.7e-100;
Matches 1162; Conservative 0; Mismatches 905; Indels 39; Gaps 7;

QY	652	ACTACTACGACGAACGAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAT	711
Db	1578	AATACCACTACACAGACGACGCTTCAACAAATGAACACCTCAACCGACGGCAATTTAAA	1637
QY	712	CGCGAGGAAATTAGTGAATCAAAACAAGTAATGAACAGCACTTTAATGATCTAAATACAGTA	771
Db	1638	AATCAAGCAACTGCTGCAAAAATGCAAGATCAAACTGTTCTCAAGAAACAAATTCCTCA	1697
QY	772	TCATCTGTAATTCACCTCCTCAAAATTCACAAATCGGAAAAATGTTTCAACAAACGCAAGAT	831
Db	1698	GTAGATAATAAAAACAACGAATGATGTAAATAGCATAGCAACAAACAGTGAAGCTTTAAAAAT	1757
QY	832	ACTTCAACTGAAGCAACACCTTCAAAACAATGAATCAGCTCCACAGAGTACAGATGCAAGT	891
Db	1758	TCTCAAAACATTAGATTTTACCACAAATCATCACCAAAACGATTTTCCCAATGCGCAAGGAACT	1817
QY	892	AATAAAGATGTAGTTAATCAAGCGGTTAATAACAAGTCGCGCTAGAAATGAGAGCACTTTAGT	951
Db	1818	AGTAAACCAAGTTTAGAACGAGAGCTGTACGTAAGTTTACGTTGCTGAACCGGTAGTA	1877
QY	952	TTAGCGGAGTAGTGCAGATGCAACCGGAGCTGGCACAGATATTAAGAATCAGTTGCAAG	1011
Db	1878	AATGCTGCTGATGCTAAAGGTACAAATGTAATGATAAAGTTACGGCAAGTAATTTCAAG	1937
QY	1012	AATGTGACAGTTGGTATTTGACTCTGGTAGCTGCTGATCGCACCAAGCAGAGTTATGTC	1071
Db	1938	TTGAAAGACTACATTTGACCCCTAATCAAAAGTGGTAACACATTTATGGCGGCAAAATTTT	1997
QY	1072	AAACTGAATATATGGTTTTTTCAGTCGCCCTAAATCTGCTGTTAAAGGTGACACATTCAAAATA	1131
Db	1998	ACAGTGACAGATAAAGTGAATCAGGGATTTATTTTACAGCGAAGTTTACAGATAGTTTAA	2057
QY	1132	ACTGTACTAAAGAAATTAACCTTAATGGTGTAACCTCAACTGCTCAAGTGCACCAACT	1191
Db	2058	ACTGGTAATGGAGACGTGGATTAATCTTAATTCAAATAATACGATGCCAAATTGCAGACATT	2117
QY	1192	ATGGCTGGAGATCNAAGTATTGGCAAAATGGTGTAATCGATAGTAGTGTAATGTTATTAT	1251
Db	2118	AAAAGTACGAATGGCGATGTTGTAGCTTAAAGCAACATATGATATCTTGACTTAAGACCGTAT	2177
QY	1252	ACA-----TTTACAGACTATGTAATACTAAAGATGATGTAAGAGCAAACTTTCACCATG	1305
Db	2178	ACATTTGCTTTACAGATTATGTAATAATAAAGAAATATTAACGGCAAAATTTTCATTA	2237
QY	1306	CCGCTTATATGTACCTCGAAAAATGTTTAAAAAGACAGGTAAATGACATTTGGCTTACTGGC	1365
Db	2238	CCTTTATTTTACAGACCGCAGCAAGGCACCTTAAATCAGAAACATATGATGCGAATATTAAAT	2297
QY	1366	ATAGTAGTACAAACAGCAACAAAAACA-----GTATTAGTAGATATTAAGAAAAATATCGTA	1420
Db	2298	ATTGCGGATGAAATGTTTAATAATAAAAAATTAATTAACATATAGTTTCGCCAAATTTGCGAGGA	2357
QY	1421	AGTTTTTA-----TAACTTATCTATTAAGGTACAAATGACCAAAATCGATAAAA	1467
Db	2358	ATTGATAAAACCAATGCGCGCAACATTTCTTCTCAATTAATTTGGTGTAGATACAGCTTCA	2417
QY	1468	ACAAATAATACGTATCGTACAGCAAAATTTATGTCAATCCAA-----GTGAGAGATAACGTTATT	1524
Db	2418	GGTCAAAACACATACAGCAACACAGTATTTGTTAAACCTTAAGCAACAGATTTTAGGTAAT	2477
QY	1525	GCGCCGT-----TTTAAACAGGTAAATTTAAACCAAAATACGGNATAGTAAATGCATTAATAGAT	1581
Db	2478	ACGTGGGTGTATATTAAGGCTCCACAGATAAAAATCGAAGAAAGATGCGGTAAGTAGT	2537
QY	1582	CAGCAAAATACAAGTATTTAAAGTATATAAGTAGATAATCAGCTGATTTATCTGAAAGT	1641
Db	2538	GTACAGATACAAAACCTCAGAAATTTTGAAGTGAAATGATACATCTAAATATACAGATAGC	2597
QY	1642	TACTTTGTGAATCCAGAAAACTTTGAGGATGTCACTAATAGTGTGAATATACATATCCCA	1701

2598	Db	TACTATGCGAGATCCAAATGACTCTAAACCTTTAAAGAAAGTAAACAGACC	2655
1702	Qy	AATCCAAATCAATATATAAGTAGAGTTTAATACGCCTGATGATCAAAATTA	1761
2658	Db	ATCTATTATGAGCATCCAAATGTAGCTAGTATTAAATTTGGTGATATTCT	2717
1762	Qy	ATAGTAGTCTTTAAATGGTCAATTATGTATCCGAATAGCAAAAGGTGATT	1821
2718	Db	GTAGTATTAGTAAGAGGCAATTACGA-CAATACAGGTAAAGAACTTAA	2776
1822	Qy	ACTTTATATCGGTATAACTCGAATATAAATTTGGCGTCTATGTCA	1881
2777	Db	TCAGAAATATGTTGATCCTCTTACAAATAGAGACTACAGTATTTTC	2836
1882	Qy	GCATTTTAATAACGGATCAGGTTCTGGTGCAGGTATCGATAAACCAG	1941
2837	Db	GAATGTGTAC-----GTTATGGTGGGAAGTCTGATGGTGATT	2888
1942	Qy	CTTGATGAGCCTGGTGAATTTGAACCAATTCAGAGGATTTCAGATT	2001
2889	Db	CCGAAAGACCCAATCCAGGGCCGCGGTTGACCCAGAACCAAGTCC	2948
2002	Qy	GATTCCTGGCAGCATCTTAATTCAGATAGCGTTTCAGATTCGGGT	2061
2949	Db	GAACCAACGCCGATCCAGAACCAAGTCCAGACCAGAACCGGAAC	3008
2062	Qy	GATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGCAAGT	2121
3009	Db	GATCCGGATTCGGATTCAGACAGTGACTCAGGCTCAGACGCACT	3068
2122	Qy	GATTCAGATTCAGACAGCGAATTCGCACTCAGCAGCGAATTC	2181
3069	Db	GACTCGAATTCAGATAGCGATTCGGATTCAGACAGTGATTCAG	3128
2182	Qy	GATTCAGATAGCGATTCGACTCAGACAGTGACTCAGATTC	2241
3129	Db	GAATTCAGATAGCGACTCAGAAATCAGATAGTGAGTTCAGATTC	3188
2242	Qy	GATAGCGATTCGACTCAGACAGTGACTCAGATTCAGATAGCG	2301
3189	Db	GACAGTGATTCAGACTCAGATAGCGATTCAGACTCAGATAGCG	3248
2302	Qy	GATTCAGATTCGACAGAGTGATTCGCACTCAGACAGCGAATTC	2361
3249	Db	GATTCAGATTCAGACAGCGACTCAGATTCAGACAGCGACT	3308
2362	Qy	GACTCAGACAGCGATTCAGATTCGCAAGTGATTCGCACTCAG	2421
3309	Db	GACTCAGACAGCGACTCAGATTCAGATAGCGATTCAGACT	3368
2422	Qy	GATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAG	2481
3369	Db	GACAGCGACTCAGACTCAGATAGCGACTCAGATTCAGATAG	3428
2482	Qy	GATTCAGATTCGACAGAGTGACTCAGATTCGCAAGTGACT	2541
3429	Db	GACTCAGATTCAGATAGCGATTCGCACTCAGACAGCGATTC	3488
2542	Qy	GATTCGACAGTGACTCAGATTCGCAAGTGACTCAGACT	2601
3489	Db	GACTCGGATAGCGATTCAGATTCAGATAGCGATTCGGAATTC	3548
2602	Qy	GCAGGTGATTCGGAATTCAGATAGTGATTCGCACTCCGACAG	2661
3549	Db	GACAGCGACTCAGACTCGGATAGCGACTCAGATTCAGACAG	3608
2662	Qy	GACTCAGACTTCGATAGCGACTCGGATTCAGATAGCGATTC	2721
3609	Db	GACTCAGACTTCGATAGCGACTCGGATTCAGATAGCGACT	3668
2722	Qy	GAATCA 2127	
3669	Db	GAATCA 3674	

RESULT 11

US-08-781-986A-96
; Sequence 96, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11050 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-96

Query Match 14.6%; Score 510; DB 4; Length 11050;
Best Local Similarity 55.2%; Pred. No. 2.7e-100;
Matches 1162; Conservative 0; Mismatches 905; Indels 39; Gaps 7;

QY	652	ACTACTAGCAACGAATCAAGCTAATACACCGGCAACAACCTCAATCAAGCAATACAAAT	711
DB	1578	AATACCCTACACAGAGCCAGCTTCAACAAATGAACACCTCAACCCGCGCAATTAA	1637
QY	712	CGGAGGAATAGTGAATCAACAGTAATGAACGACTTTAATGATCTAATACAGTA	771
DB	1638	AATCAAGCAACTGTGCAAAATCAAGATCAAACTGTTCTCAAGAGCAAAATTTCA	1697
QY	772	TCATCTGTAATTCACCTCAAAATTTCTCAAAATGCGGAAATGTTTCAACACGCAAGT	831
DB	1698	GTAGATATAAACAACGAATGATCTAATAGCATAGCAACAACAGTGAGCTTAAAT	1757
QY	832	ACTTCAACTGAAGCAACACCTTTCAACAAATGAATCAAGCTCCACAGATACAGATGCAAGT	891
DB	1758	TCTCAACATTTAGATTTTACCACATCATCACCAACAAACGATTTCCAAATGCGCAAGAACT	1817
QY	892	AATGAAGATGTAGTTAATCAAGCGTTAATCAACAGTGCGCTAGATGAGAGCATTTAGT	951
DB	1818	AGTAACCAAGTGTAGAACGAGAGCTGTACGTAGTTTAGCTTGTGCTGAACCGGTAGTA	1877
QY	952	TTAGCGGAGTAGCTGCAGATGCACCGGAGCTGGCAACAGATATTACGAATCAGTTGACG	1011
DB	1878	AATGCTGCTGATGCTAAAGGTACAATGTAAATGATAAAGTTACGGCAAGTAATTTCAAG	1937
QY	1012	AATGTACAGTTGGTATTGACTCTCGTAGACTGTGTATCCGCCAACCAAGAGTTATGTC	1071

DB	1938	TTAGAAAAGACTACATTTGACCCCTAATCAAAGTGGTAACACATTTATGGCGCAAAATTT	1997
QY	1072	AAACTGAATTATGGTTTTCAGTGCCCTAATTTCTGCTGTTAAAGGTGACACATTCAAATA	1131
DB	1998	ACAGTGACAGATAAAGTGAATCAGGGGATTTTACAGCGAAGTTTACAGATAGTTTA	2057
QY	1132	ACTGTACCTAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGTCTAAAGTGCACCAATT	1191
DB	2058	ACTGGTAATCGAGACGTGGATTATTCTPAATTCBAATATACGATGCCAATTCGACACATT	2117
QY	1192	ATGGCTGGAGATCAAGTATTCGAAAATGGGTAAATCGATAGTGGTAATGTTTATTTAT	1251
DB	2118	AAAAGTACGAATGGCGATGTTGTAGCTAAAGCAACATATGATATCTTGACTAAGACGAT	2177
QY	1252	ACA-----TTTACAGACTATGTAATACTAAAGATGATGTAAGAGCAACTTTGACCATG	1305
DB	2178	ACATTTGCTTTACAGATTATGTAATAATAAAGAAATAATTAACGGCAATTTTCATTA	2237
QY	1306	CCCGCTTATATTGACCTTGAAAATGTTAAAGAGACAGGTAATGTGACATTTGGCTACTGCG	1365
DB	2238	CCTTTATTTACAGACCGAGCAAGGCCTTAATCAGGAACATATGATGCGAATTAAT	2297
QY	1366	ATAGTAGTACAACAGCAAAACAAACA-----GTATTAGTAGATTATGAAAAATATGTA	1420
DB	2298	ATTGGGATGAATGTTTAAATAAAAAATTACTTATTAATACTAGTTCGCCAATTCGAGGA	2357
QY	1421	AGTTTTA-----TAACTTATCTATTAAAGGTACAATTGACCAATCGATAA	1467
DB	2358	ATTGATAAACCAATGGCGCGAACATTTCTTCTCAAAATATTGGTGTAGATACAGCTTCA	2417
QY	1468	ACAAATAATACGTATCGTCAGACAAATTTATGTCATCCAA---GTGGAGATAAACGTTATT	1524
DB	2418	GGTCAAAACACATACAGCAACACATATTTCTTAACCCCTAAGCAACGAGTTTAGGTAAT	2477
QY	1525	GGCGCGGT---TTTAAACAGGTAATTTAAACCAATACGGATAGTAATGCAATTAATAGAT	1581
DB	2478	ACGTGGGTGTATATTAAAGGCTACCAAGATAAAAATCGAAGAAAGTAGCGGTAAAGTAAGT	2537
QY	1582	CAGCAAAATACAAGTATTAAAGTATATAAAGTATATAAGTATGAGCTGATTTATCTGAAAAGT	1641
DB	2538	GCTACAGATCAAAACCTGAGAAATTTTGAAGTGAATGATACATCTAAATTTACAGATAGC	2597
QY	1642	TACTTTGTGAATCCAGAAAACCTTTGAGGATGTCATATAGTGTGAATATTACATTCCTCA	1701
DB	2598	TACTATGCAGATCCAAATGACTCTTAACCTTAAAGAAAGTAAACAGACCAATTTAAAAATAGA	2657
QY	1702	AATCCAAATCAATATAAAGTAGAGTTTAAATAGCGCTGATGATCAAAATTTACAAACCCGTAT	1761
DB	2658	ATCTATTATGAGCATCCAAATGTAGTAGTATTAAATTTGGTGATATTACTTAAACATAT	2717
QY	1762	ATAGTAGTTGTTAATGGTTCATATTGATCCGAATAGCAAGGTGATTTAGCTTTAGCTTCA	1821
DB	2718	GTAGTATTAGTAGAAGGGCAITTAGA-CAATACAGGTGAAGAACTTAAAACTCAGGTTAT	2776
QY	1822	ACTTTATATGGGTATACTCGAAATATAATTTGGCGCTCTATGTATGCGGACCAACGAAGTA	1881
DB	2777	TCAAGAAAATCTTGATCTCTGTAAACAAATAGAGACTACAGTATTTTCGGTTTGAATAATGA	2836
QY	1882	GCATTTAATACGGATCAGGTTCTGGTGACGGTATCGATAAACAGGTTGTTCTCTGAACAA	1941
DB	2837	GAATGTTGTAC-----GTTATGGTGGTGAAGTGTCTGATGGTGAATTCAGCAGTAAT	2888
QY	1942	CCTGATGAGCTGTGGAATTTGAACCAATTCAGAGGAATTCAGATTCCTGACCCAGGTTCA	2001
DB	2889	CCGAAGACCAACTCCAGGCGCCCGGTTGACCCAGNACCAAGTCCAGACCCAGAACCA	2948
QY	2002	GATTCTGCAGCGATTCTTAATTCAGATAGCGGTTTCAAGATTCGGGTAGTGTATTCATCA	2061
DB	2949	GAAACCAAGCCAGATCCAGAACCAAGTCCAGACCCAGAACCCGGAACCAAGCCAGACCGG	3008
QY	2062	GATAGTGGTTTCAAGATTCAGCGAGTGTATTCAGATTCAGCAAGTGTATTCAGATCTCAGCGAGT	2121

Db 3009 GATCCGGATTCCGATTTCAGACAGTGAAGTCTCAGGCTCAGACAGCGACTCAGGTTTCAGATAGC 3068
QY 2122 GATTTCAGATTTCAGACAGGATTCGAGTTCAGGAGCGATTCGAGTTCAGACAGTGAAGTCTCG 2181
Db 3069 GACTTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 3128
QY 2182 GATTTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 2241
Db 3129 GAATTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 3188
QY 2242 GATTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 2301
Db 3189 GACAGTGAATTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 3248
QY 2302 GATTTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 2361
Db 3249 GATTTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 3308
QY 2362 GACTTCAGACAGCGGATTCAGATTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2421
Db 3309 GACTTCAGACAGCGGATTCAGATTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 3368
QY 2422 GATTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 2481
Db 3369 GACAGCGGATTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 3428
QY 2482 GATTTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 2541
Db 3429 GACTTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 3488
QY 2542 GATTTCGAGTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 2601
Db 3489 GACTTCGAGTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 3548
QY 2602 GCGAGTGAATTCGAGTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 2661
Db 3549 GACAGCGGATTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 3608
QY 2662 GACTTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 2721
Db 3609 GACTTCAGATTTCAGATAGCGATTCGAGTTCAGACAGTGAAGTCTCAGGAGCGGACTCA 3668
QY 2722 GAATCA 2727
Db 3669 GATTCA 3674

RESULT 12

US-09-147-405B-14
; Sequence 14, Application US/09147405B
; Patent No. 6733759
; GENERAL INFORMATION:
; APPLICANT: Guss, Bengt
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Plock, Jan-Ingar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; TITLE OF INVENTION: Coagulase-Negative Staphylococcus
; FILE REFERENCE: guss 09/147405
; CURRENT APPLICATION NUMBER: US/09147405B
; CURRENT FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 3600
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:

; NAME/KEY: CDS
; LOCATION: (33)..(3308)
US-09-147-405B-14
Query Match 13.2%; Score 462; DB 4; Length 3600;
Best Local Similarity 75.6%; Pred. No. 3.9e-90;
Matches 573; Conservative 0; Mismatches 185; Indels 0; Gaps 0;
QY 1984 GATTCGACCCAGGTTTCAGATTCTCGCAGCGGATTCATATTCAGATAGCGGTTTCAGATTTCG 2043
Db 2412 GATTCGACAAACAGATGCTGATGGGAAGAGTTTCATGTAACTACTGATCATGATGAC 2471
QY 2044 GGTAGTGAATTCATACATCATAGTAGTGGTTTCAGATTTCAGCGAGTGAATTCAGATTTCAG 2103
Db 2472 TTATGATAGATAACCGGATACATGATGACGAATCGGATTCGATAGTACTCAGACAGC 2531
QY 2104 GATTCAGATTCAGCGAGTGAATTCAGATTTCAGCGAGGATTCGAGTTCAGCGAGGATTC 2163
Db 2532 GACTTCAGATTTCGATAGTGAATTCAGATTTCAGCGAGTTCGAGTTCAGCGAGGATTC 2591
QY 2164 GACTTCAGACAATGACTTCGAGTTCAGATAGCGGATTCGATTCAGACAGTGAATTCAGATTTC 2223
Db 2592 GATTCAGACAGCGGATTCGAGTTCGATAGCGGATTCGATTCAGCGAGTTCGAGTTCAGCG 2651
QY 2224 GACAGTGAATTCAGATTTCAGATAGCGGATTCGATTCAGACAGTGAATTCAGATTTCAG 2283
Db 2652 GACAGTGAATTCAGATTTCAGACAGCGGATTCGATTCGATAGTGAATTCAGATTTCAG 2711
QY 2284 GATTCAGATTTCAGATAGCGGATTCAGATTTCGAGTGAATTCGAGTTCGAGTTCAGCGAG 2343
Db 2712 GACTTCAGATTTCGATAGTGAATTCAGATTTCAGCGAGTTCGAGTTCAGCGAGGATTC 2771
QY 2344 GACTTCGAGCGAGTTCGAGTTCGAGTTCAGCGAGGATTCGAGTTCGAGCGAGTTCGAGTTC 2403
Db 2772 GATTCGAGTGAATTCAGATTTCAGCGAGGATTCGAGTTCGAGTTCGAGTTCGAGTTCAG 2831
QY 2404 GATTCGAGTTCGAGTTCAGATAGCGGATTCAGATTTCAGACAGCGGATTCAGATTTCAG 2463
Db 2832 GACAGCGGATTCAGATTTCGATAGTGAATTCAGATTTCAGCGAGTTCAGATTTCAGCG 2891
QY 2464 GATTCAGATTTCAGATAGCGGATTCAGATTTCGAGTGAATTCGAGTTCGAGTTCGAGTTC 2523
Db 2892 GATTCAGATTTCGATAGTGAATTCAGATTTCAGCGAGTTCGAGTTCGAGTTCGAGTTCAG 2951
QY 2524 GATTCAGATAGCGGATTCAGATTTCGAGTGAATTCGAGTTCGAGTTCGAGTTCGAGTTCAG 2583
Db 2952 GATTCGATAGTGAATTCAGATTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 3011
QY 2584 GACAGTGAATTCGAGTTCAGCGAGTGAATTCGAGTTCAGATAGTGAATTCGAGTTCGAG 2643
Db 3012 GATTCGATTCAGATTTCAGATTTCAGCGAGGATTCAGATTTCGATAGTGAATTCAGATTTC 3071
QY 2644 GACTTCGAGTTCAGATAGCGGATTCAGATTTCGAGTGAATTCGAGTTCGAGTTCGAGTTCAG 2703
Db 3072 GATTCAGATTTCGATAGTGAATTCAGATTTCAGCGAGTTCGAGTTCGAGTTCGAGTTCAG 3131
QY 2704 GACTTCAGATAGCGGATTCAGATTTCAGCGAGGATTCAGATTTCAGATTTCAGATTTCAG 2741
Db 3132 GACTTCAGACAGCGATTCAGATTTCAGCGAGGATTCAGATTTCAGATTTCAGATTTCAG 3169

RESULT 13

US-08-956-171E-454/C
; Sequence 454, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences


```

Best Local Similarity 71.4%; Pred. No. 3.5e-87;
Matches 641; Conservative 0; Mismatches 236; Indels 21; Gaps 3;
1865 CATGGGACACGAGTAGACATTTAAACGAGTACAGGTTCTGGTGACGGTATCGATAAAC 1924
1866 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2374 CAACTGARACAGATGAAATGGTAAATACCGCTTTGATATTTAGATAGTGGTAAATACA 2315
2375 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1925 CAGTTGTTCTGAACACCTGATGAGCTGGTGAATGAACCAATCCAGAGGATTCAG 1984
1926 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2314 AGTTATCTTTGAAACCTGCTGGCTTAACCTAAACAGGTACAAATACAACTGAAGATG 2255
2315 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1985 ATTCTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTTCAGATTCGG 2044
1986 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2254 AATAAGATCGGATGGTGGGAGATTGATGAACAAATACGGATCATGAT--GATTTCA 2198
2255 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2045 GTAGTGAATCTACATCAGATAGTGGTTTCAGATTCAGGAGTGAATTCAGATTCAGCAAGTG 2104
2046 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2197 CACTTGATAATGCTACTACGAAGAAGAAACATCAGATAGCGACTCAGATTTCTGACAGCG 2138
2198 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2105 ATTCAGACTCAGCGAGTGTGATTCAGATTCAGCAAGCGATTCGAGCTCAGCGAGCGATTCGG 2164
2106 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2137 ATTCAGACTCAGATAGCGACTCAGATTCAGATTCAGATAGCGACTCAGATTCAGACAGCGAT-- 2082
2138 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2165 ACTCAGCAATGACTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCGG 2224
2166 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2081 --TCAGACAGCGACTCAGACTCAGATAGCGATTCAGATTCAGACAGCGACTCAGACTCAG 2024
2082 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2225 ACAGTGAATCAGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGATAGCG 2284
2226 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2023 ACAGCGATTCAGACTCGGATTCAGATAGCGACTCAGACTCAGATAGCGACTCAGATTCGGATAGCG 1964
2024 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2285 ATTCAGATTCAGATAGCGATTCAGATTCAGACAGTGAATTCGAGCTCAGACAGCGATTCG 2344
2286 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1963 ACTCAGACTCAGATAGCGATTCAGATTCAGATAGCGATTCGAGCTCAGACAGTGAATTC-- 1906
1964 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2345 ACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGAGTGAATTCGAGTTCAG 2404
2346 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1905 -----AGATTCAGACTCAGATAGCGACTCAGATTCGAGCGGATTCAGACTCAG 1856
1906 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2405 ATAGCGATTCGAGTTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCG 2464
2406 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1855 ACAGCGACTCAGACTCAGACAGTGAATTCAGATTCAGACAGCGACTCAGATTCAGATTCAGATAGCG 1796
1856 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2465 ATTCAGATTCAGATAGCGATTCAGATTCGACAGTGAATTCGAGTTCGAGTGAATTCGAGTTCG 2524
2466 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1795 ACTCAGACTCAGATAGCGACTCAGATTCAGATAGCGATTCGAGCTCAGACAGCGACTCAG 1736
1796 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2525 ATTCAGATAGCGATTCAGATTCGACAGTGAATTCAGATTCGAGTGAATTCGAGTTCGAGTTCG 2584
2526 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1735 ATTCAGATAGCGATTCAGATTCAGATAGCGACTCAGATTCGAGCGGATTCAGACTCAG 1676
1736 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2585 ACAGTGAATTCGAGTTCAGCGAGTGAATTCGAGTTCAGATAGCGACTCAGACTCAGAGTG 2644
2586 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1675 ATAGCGATTCAGACTCAGACAGCGATTCAGATTCAGATAGCGACTCAGACTCAGATAGCG 1616
1676 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2645 ACTCGGATTCAGATAGCGACTCAGACTCAGATAGCGACTCAGATTCAGATAGCGATTCG 2704
2646 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1615 ACTCAGACTCAGATAGCGATTCAGATTCAGACAGCGACTCAGATTCAGATAGCGATTCG 1556
1616 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2705 ACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGATTC 2762
2706 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1555 ACTCAGACAGCGACTCAGATTCAGATAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1498
1556 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 15
US-09-710-279-31/c
; Sequence 31, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279

```

```

; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-31

```

```

Query Match 12.5%; Score 436.8; DB 4; Length 756;
Best Local Similarity 78.5%; Pred. No. 5.9e-85;
Matches 567; Conservative 0; Mismatches 137; Indels 18; Gaps 3;
QY 2074 GATTCAGCGAGTGAATTCAGATTCAGCAAGTGAATTCAGACTCAGCGAGTGAATTCAGATTCA 2133
DB 727 GATTCAGACAGTGAATTCAGACGAGATAGTGAATTCAGACTCAGCGAGTGAATTCAGATTCA 668
QY 2134 GCAAGCGATTCGAGTTCAGCGAGCGATTCGAGTTCAGCAATGAGTTCGAGTTCAGATTCA 2193
DB 667 GACAGCGATTCGAGCGAGCGAGTGAATTCAGATTCAGATAGTGAATTCAGACTCAGACAGC 608
QY 2194 GATTCGAGTTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2247
DB 607 GATTCGAGTTCGAGTGAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 548
QY 2248 GATTCGAGTTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2307
DB 547 GATTCGAGCGAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 488
QY 2308 GATTCGAGCGAGTTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2361
DB 487 GATTCGAGTGAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 428
QY 2362 GATTCGAGCGAGTTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2421
DB 427 GACGCGAGCGAGTGAATTCAGATTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 368
QY 2422 GATAGCGACTCAGATTCAGACAGCGAGTTCAGATTCAGACAGCGAGTTCAGATTCAGATTCAG 2481
DB 367 GATAGTGAATTCGAGTTCGAGCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 308
QY 2482 GATTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2541
DB 307 GATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 248
QY 2542 GATTCGCGAGTGAATTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2601
DB 247 GACGCGAGCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 188
QY 2602 GCGAGTGAATTCGAGTTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2655
DB 187 GACAGTGAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 128
QY 2656 GATAGCGACTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2715
DB 127 GACAGTGAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 68
QY 2716 GATTCAGATTCAGACAGCGAGTTCAGATTCAGACAGCGAGTTCAGATTCAGACAGCGAGTTCAG 2775
DB 67 GATTCGAGTTCAGATAGCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAGTCA 8
QY 2776 GA 2777
DB 7 GA 6

```

```

Search completed: November 5, 2004, 12:59:33
Job time : 285 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2004, 06:05:04 ; Search time 117 Seconds
(without alignments)
6924.920 Million cell updates/sec

Title: US-09-679-643-1

Perfect score: 3498

Sequence: 1 ggtaccataattacacatc.....gaagagtataagaagagctt 3498

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 693957 seqs, 115811272 residues

Total number of hits satisfying chosen parameters: 1387914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1592.6	45.5	1651	8	US-60-615-573-18588
2	1393.6	39.8	1600	8	US-60-615-573-18330
3	217.8	6.2	404	8	US-60-615-573-14144
4	169.4	4.8	180	8	US-60-615-573-18462
5	141.4	4.0	4050	6	US-60-615-573-14132
6	99	2.8	171	8	US-60-615-573-14147
7	85	2.4	98412	6	US-60-615-573-18342
8	79.8	2.3	2701	8	US-60-615-573-18342
9	79.4	2.3	388	6	US-60-615-573-1506
10	73.8	2.1	427	6	US-60-615-573-14147
11	72.6	2.1	2493	8	US-60-615-573-14147
12	71.6	2.0	5192	8	US-60-615-573-14144
13	71.2	2.0	404	8	US-60-615-573-14144
14	67.2	1.9	105306	6	US-60-615-573-18593
15	67	1.9	2032	8	US-60-615-573-18593
16	66.8	1.9	337	8	US-60-615-573-14146
17	65.6	1.9	5469	6	US-60-615-573-18596
18	64.4	1.8	1601	8	US-60-615-573-18596
19	64	1.8	47475	6	US-60-615-573-18596
20	64	1.8	2242716	6	US-10-915-740A-13
21	61.8	1.8	74700	1	PCT-US03-40884-3
22	60.4	1.7	193	8	US-60-615-573-17864
23	60	1.7	129757	6	US-60-615-573-17864
24	57	1.6	384	6	US-10-746-294A-88
25	55.2	1.6	863	6	US-10-220-366A-2561
					US-10-479-081-32

c	26	54.6	1.6	7429	6	US-10-873-332-61	Sequence 61, Appl
c	27	54.6	1.6	256145	6	US-10-856-118-1	Sequence 1, Appl
c	28	54.2	1.5	71184	6	US-10-746-294A-69	Sequence 69, Appl
	29	53.8	1.5	3893	6	US-10-956-160-5849	Sequence 5849, Ap
c	30	53.4	1.5	25138	6	US-10-832-622B-18	Sequence 18, Appl
c	31	53.2	1.5	2100	8	US-60-615-573-18214	Sequence 18214, A
c	32	52.6	1.5	7001	6	US-10-172-086A-60	Sequence 60, Appl
c	33	52.2	1.5	759	8	US-60-615-573-14092	Sequence 14092, A
	34	50.6	1.4	123805	6	US-10-849-551-3	Sequence 3, Appl
	35	50.4	1.4	384	6	US-10-948-737-10577	Sequence 10577, A
	36	50.4	1.4	15075	1	PCT-US04-18902-16	Sequence 16, Appl
	37	50.2	1.4	2676	8	US-60-615-573-18495	Sequence 18495, A
	38	49.4	1.4	88010	6	US-10-746-294A-78	Sequence 78, Appl
	39	49.4	1.4	110980	6	US-10-746-294A-80	Sequence 80, Appl
	40	49.4	1.4	123805	6	US-10-849-551-3	Sequence 3, Appl
	41	49.2	1.4	2511	6	US-10-956-157-4179	Sequence 1781, Ap
	42	49.2	1.4	2821	6	US-10-956-157-1781	Sequence 1781, Ap
	43	49	1.4	7001	6	US-10-172-086A-59	Sequence 59, Appl
	44	49	1.4	78921	6	US-10-746-294A-65	Sequence 65, Appl
	45	49	1.4	157000	6	US-10-746-294A-121	Sequence 121, App

ALIGNMENTS

RESULT 1
US-60-615-573-18588

; Sequence 18588, Application US/60615573

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Murphs, William M

; APPLICANT: Murphy, Ellen

; APPLICANT: Olmsted, Stephen

; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES

; FILE REFERENCE: 031896-084099 (AM 101724)

; CURRENT APPLICATION NUMBER: US/60/615,573

; CURRENT FILING DATE: 2004-10-05

; NUMBER OF SEQ ID NOS: 18598

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 18588

; LENGTH: 1651

; TYPE: DNA

; ORGANISM: Staphylococcus aureus

US-60-615-573-18588

Query Match 45.5%; Score 1592.6; DB 8; Length 1651;

Best Local Similarity 93.2%; Pred. No. 0;

Matches 1538; Conservative 111; Mismatches 1; Indels 1; Gaps 1;

QY	301	ATGAATATGAGNAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA	360
Db	1	ATGAATATGAGNAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA	60
QY	361	GTGCTTGTAGGTACGTTAATCGGTTTGGACTACTCAGCAGTAAAGAACGAGATGCAAGT	420
Db	61	GTGCTTGTAGGTACGTTAATCGGTTTGGACTACTCAGCAGTAAAGAACGAGATGCAAGT	120
QY	421	GAATAATGTTTACGCAATCTGTATAGCGCAAGTAACGAAAGCAAAAGTAATGATTTCAAGT	480
Db	121	GAATAATGTTTACGCAATCTGTATAGCGCAAGTAACGAAAGCAAAAGTAATGATTTCAAGT	180
QY	481	AGCGTTAGTGTCTGCACCTAAAAACAGACGACACAAAGTGTAGTACTAAACATCGTCA	540
Db	181	AGCGTTAGTGTCTGCACCTAAAAACAGACGACACAAAGTGTAGTACTAAACATCGTCA	240
QY	541	AACACTAATAATGGGAAACGAGTGTGGCGGCAAAATCCAGACACACAGAGAAACGACACA	600
Db	241	AACACTAATAATGGGAAACGAGTGTGGCGGCAAAATCCAGACACACAGAGAAACGACACA	300
QY	601	TCATCATCAACAAATGCAACTACGGAAGAAACCCCGGTAACTGGTGAAGCTACTACTACG	660
Db	301	TCATCATCAACAAATGCAACTACGGAAGAAACCCCGGTAACTGGTGAAGCTACTACTACG	360

```
QY 661 ACACGATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA 720
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 RCAAGCAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA 420
QY 721 TTAGTGAATCAAAACAAAGTAATGAACGACCTTTTAAATGATACATAATACAGTATCATCTGTA 780
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 TYAGTGAATCAAAACAAAGTAATGAACGACCTTTTAAATGATACATAATACAGTATCATCTGTA 480
QY 781 AATTCACCTCAAAATTTCTAAGTTCGCGGAAATGTTTCAACGACGAGATCTCAACT 840
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 AATTCACCTCAAAATTTCTAAGTTCGCGGAAATGTTTCAACGACGAGATCTCAACT 540
QY 841 GAAGCAACACCTTCAAAACAAAGTAATGAACGACCTTCAACGAGTACAGATGAAGTAATAAGAT 900
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
541 GAAGCAACACCTTCAAAACAAAGTAATGAACGACCTTCAACGAGTACAGATGAAGTAATAAGAT 600
QY 901 GTAGTGAATCAAGCGGTTTAAATACAGTTCGCTGATAGATGAGAGCAATTTAGTTTAGCGCA 960
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
601 GTAGTGAATCAAGCGGTTTAAATACAGTTCGCTGATAGATGAGAGCAATTTAGTTTAGCGGCA 660
QY 961 GTAGTGCAGATGACACCGGAGCTGGACAGATATTACGAATCAGTTGACGAATGTGACA 1020
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 GTAGTGCAGATGACACCGGAGCTGGACAGATATTACGAATCAGTTGACGAATGTGACA 720
QY 1021 GTTGGTATTGACTCTGGTACGACTGTGATCGCACCAAGCAGGTTATGTCAAACTGAAT 1080
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
721 GTTGGTATTGACTCTGGTACGACTGTGATCGCACCAAGCAGGTTATGTCAAACTGAAT 780
QY 1081 TATGGTATTTTCAAGTTCGCTGTTTAAAGTGCACATTCACAAATCAACTGACT 1140
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
781 TATGGTATTTTCAAGTTCGCTGTTTAAAGTGCACATTCACAAATCAACTGACT 840
QY 1141 AAGATTAATCAAGTTCGCTGTTTAAAGTGCACATTCACAAATCAACTGACT 1200
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
841 AAGATTAATCAAGTTCGCTGTTTAAAGTGCACATTCACAAATCAACTGACT 900
QY 1201 GATCAAGTATTGGCAAAATGTTTAAATGATGATGTTTAAATGATGATGTTTAAATGATGATGAT 1260
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
901 GATCAAGTATTGGCAAAATGTTTAAATGATGATGTTTAAATGATGATGTTTAAATGATGATGAT 960
QY 1261 GACTATGTAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
961 GACTATGTAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1321 CCTGAAATGTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1021 CCTGAAATGTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1381 GCACAAACAAACAGTATTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1081 GCACAAACAAACAGTATTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1441 AAAGGTACAAATGACCAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1141 AAAGGTACAAATGACCAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1501 AATCCAAAGTGGAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1201 AATCCAAAGTGGAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1561 GATAGTAATGCAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1261 GATAGTAATGCAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1620 TGCAGCTGATTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1321 TGCACGAGATTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1680 TAGTGTGAATTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1381 TAGTGTGAATTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1740 TGATCAAAATTACACACCGTATATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799
```

```
Db 1441 TGATCAAAATWACACACCTATATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1800 AGGTGATTAGCTTTAGCTTTCAACTTTTATATGGGTATAAATCGGAATATAAATTTGCGGCTC 1859
Db 1501 AGGTGATTAGCTTTAGCTTTCAACTTTTATATGGGTATAAATCGGAATATAAATTTGCGGCTC 1560
QY 1860 TATGTCATGGGACACAAAGTAGCATTTTAAATACCGGATCAGGTTCTGGTGACGGTATCGA 1919
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1561 WATGTCATGGGAYAAAGTAGCATTTTAAATACCGGATCAGGTTCTGGTGACGGTATCGA 1620
QY 1920 TAAACCGATGTTCTTCTGAAACAACTGATGAG 1950
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1621 TAAACCTGTTCTTCTGAAACAACTGATGAG 1651

RESULT 2
US-60-615-573-18330
; Sequence 18330, Application US/60615573
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18330
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-60-615-573-18330

Query Match 39.8%; Score 1393.6; DB 8; Length 1600;
Best Local Similarity 91.9%; Pred. No. 2.6e-296;
Matches 1471; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 301 ATGAATATGAAGAAAGAAAGAAACACGCAATTCGGAATAATCGATTGGCGTGCTTCA 360
Db 1 ATGAATATGAAGAAAGAAAGAAACACGCAATTCGGAATAATCGATTGGCGTGCTTCA 60
QY 361 GTGCTTTAGGTAGTACCTTAAATCGTTTGGACTACTCAGCAGTAAAGAGCAGATCAACT 420
Db 61 GTGCTTTAGGTAGTACCTTAAATCGTTTGGACTACTCAGCAGTAAAGAGCAGATCAACT 120
QY 421 GAAATAGTGTTCGCAATCTGATAGCGCAAGTAAAGAAAGCAAAAGTAATGATTTCAACT 480
Db 121 GAAATAGTGTTCGCAATCTGATAGCGCAAGTAAAGAAAGCAAAAGTAATGATTTCAACT 180
QY 481 AGCGTTAGTGTGACCTTAAACACGACACAAACGTCGAGTGATCTAAACATCGTCA 540
Db 181 AGCGTTAGTGTGACCTTAAACACGACACAAACGTCGAGTGATCTAAACATCGTCA 240
QY 541 AACACTTAATAATGCGGAAACGAGTGTGCGCAAAATCCAGCACAAACAGGAAACGACAAA 600
Db 241 AACACTTAATAATGCGGAAACGAGTGTGCGCAAAATCCAGCACAAACAGGAAACGACAAA 300
QY 601 TCATCATCAACAAATGCAACTACGGAAGAAACCGCGGTAACTGGTGAAGCTACTACTAG 660
Db 301 TCATCATCAACAAATGCAACTACGGAAGAAACCGCGGTAACTGGTGAAGCTACTACTAG 360
QY 661 ACACGATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGAA 720
Db 361 ACACGATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGAA 420
QY 721 TTAGTGAATCAAAACAAAGTAATGAACGACCTTTTAAATGATACATAATACAGTATCATCTGTA 780
Db 421 TTAGTGAATCAAAACAAAGTAATGAACGACCTTTTAAATGATACATAATACAGTATCATCTGTA 480
QY 781 AATTCACCTCAAAATTTCTAAGTTCGCGGAAATGTTTCAACGACGAGATCTCAACT 840
```

Db 481 AATTCACTCAAAATCTACAAATCGGAAAATGTTTCAACAACGCAAGATACCTCAACT 540
QY
Db 841 GAAGCAACACCTTCAACAATGAATCAGCTCCACAGACTCAGATGCAAGTAATAAAGAT 900
QY
Db 541 GAAGCAACACCTTCAACAATGAATCAGCTCCACAGATCAGATGCAAGTAATAAAGAT 600
QY
Db 901 GTAGTTAATCAAGCGGTTAATACAAAGTCGCGCTAGAATGAGAGCATTTAGTTTACGCGCA 960
Db 601 GTAGTTAATCAAGCGGTTAATACAAAGTCGCGCTAGAATGAGAGCATTTAGTTTACGCGCA 660
QY
Db 961 GTAGCTCAGATGACCGGAGCTGGCCACAGATATTACGAATCAGTTGACGAATGTGACA 1020
Db 661 GTAGCTCAGATGACCGGAGCTGGCCACAGATATTACGAATCAGTTGACGAATGTGAAA 720
QY
Db 1021 GTTGGTATTGACTCTGCTGACACTGTGTATCCGACCAAGCAGGTATGTCAAACTGAAT 1080
Db 721 GTTACATTGACTCTGCTGACACTGTGTATCCGACCAAGCAGGTATGTCAAACTGAAT 780
QY
Db 1081 TATGGTTTTTCAAGCTTAATCTGCTGTTAAAGGTGACACATTCAAAATAAATCTGTACCT 1140
Db 781 TATGGTTTTTCAAGCTTAATCTGCTGTTAAAGGTGACACATTCAAAATAAATCTGTACCT 840
QY
Db 1141 AAGAAATTAACCTTAATGTTGTTAACTTCAACTCTAAAGTGCACCAATTAATGGCTGGA 1200
Db 841 AAGAAATTAACCTTAATGTTGTTAACTTCAACTCTAAAGTGCACCAATTAATGGCTGGA 900
QY
Db 1201 GATCAAGTATTGGCAATGTTGAATCGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 901 GATCAAGTATTGGCAATGTTGAATCGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY
Db 1261 GACTATGTAATACTAAAGATGATGTAAGCAATTTTGGCCCTGCTTATATTGAC 1320
Db 961 GACTATGTTGATTAAGAAATGTAACAGCTTAATTAATTAATTAATTAATTAATTAATTAAT 1020
QY
Db 1321 CCTGAAATGTTTAAAGACAGGTAATGTGACATTTGGCTACTGSCATAGGTAGTACAA 1380
Db 1021 CCTGAAATGTTTAAAGACAGGTAATGTGACATTTGGCTACTGSCATAGGTAGTACAA 1080
QY
Db 1381 GCAAAACAAACAGTATTAGTAGATTAATGAAATATGTTAAAGTATGTTTATTAATCTTCTATT 1440
Db 1081 GCTAGTAGACAGTATTATCGACTATGAGAAATATGGAACAATTCCTAATAATTTATCAAT 1140
QY
Db 1441 AAGAGTCAATTCACCAATTCGATAAACAATAATACGTATCGTCAGACAAATTTATGTC 1500
Db 1141 AAGAGTCAATTCGATAAACAATAATACGTATCGTCAGACAAATTTATGTC 1200
QY
Db 1501 AATCCAGTGGAGATAACGTTATGCGCCGGTTTAAACAGGTAAATTTAAACCAAAATACG 1560
Db 1201 AATCCAGTGGAGATAACGTTATGCGCCGGTTTAAACAGGTAAATTTAAACCAAAATACG 1260
QY
Db 1561 GATAGTAATGCATTAATAGATCAGCAAAATACAGTATTAAAGTATATAAGTAGATTAAT 1620
Db 1261 AAGAGTAATGCGTTAATAGATCAGCAAAACACGTATTAAGTATTATAGATGCGATAAT 1320
QY
Db 1621 GCAGCTGATTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTTCAGGAGTGCATTAAT 1680
Db 1321 GCTAATGATTATCTGAAAGTTATTATGUGAATCCTAGCGATTTTGAAGATGTAATTAAT 1380
QY
Db 1681 AGTGTGAATTAATACATTCCAAAATCCAAATCAATATAAAGTAGAGTTTAAATACGCTGAT 1740
Db 1381 CAAGTTTAGAATTTCAATTCCAAAATGCTAATCAATCAAAAGTAGAATTTCTCCTACGACGAT 1440
QY
Db 1741 GATCAAAATTAACAACCGGTATATAGTAGTTGTTAATGGTCATATTGATCGAATAGCAA 1800
Db 1441 GACCAAAATTAACAACCGGTATATAGTAGTTGTTAATGGCCATATTGATCCTGCTAGTACA 1500
QY
Db 1801 GGTGATTAGCTTTAGCTTCAACTTTATATGGGTATAAATCGAATATAATTTGCGCTCT 1860
Db 1501 GGTGATTAGCTTTAGCTTCAACTTTATATGGGTATAAATCGAATATAATTTATGGAGATCT 1560
QY
Db 1861 ATGTCTGGACAAACGAAGTAGCATTTAATAACGGATCAG 1900

Db 1561 ATGTCTATGGGACAAAGTAGCATTTTAATAACGGATCAG 1600

RESULT 3

US-60-615-573-14144/c
; Sequence 14144, Application US/60615573
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14144
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-60-615-573-14144

Query Match 6.2%; Score 217.8; DB 8; Length 404;

Best Local Similarity 78.5%; Pred. No. 4.2e-39;

Matches 252; Conservative 6; Mismatches 63; Indels 0; Gaps 0;

QY 2357 ATTCGAGCTCAGACAGCGGATTCAGATTCGACAGTGTATCCGACTCAGATGCGGATTCG 2416

Db 327 ATTCGAGTAAATGATTAGATACAGATATCGTTTCAATAGTACTCAGAAAAATGACACAT 268

QY 2417 ACTCAGATACGAGCTCAGATTCAGACAGCGAATTCAGATTCAGACAGCGATTCAGATTCAG 2476

Db 267 ATTTAGATAGTGAATTCAGACTCAGATAGTAGTACATTCAGATTCAGATTCAGATTCAG 208

QY 2477 ATAGCGATTCAGATTCGACAGTGAATTCAGATTCGACAGTGAATTCGATTCGATTCAGATAGCG 2536

Db 207 ATAGTATTCAGATTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 148

QY 2537 ATTCAGATTCGACAGTGAATTCAGATTCGACAGTGAATTCGACAGTGAATTCGACAGTGAATTCG 2596

Db 147 ACTCAGATTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 88

QY 2597 ATTCAGCGATGATTCGATTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 2656

Db 87 ACTCAGATGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 28

QY 2657 ATAGCGACTCAGACTCGGATA 2677

Db 27 ATAGTATTCAGACTCTGGTA 7

RESULT 4

US-60-615-573-18462/c
; Sequence 18462, Application US/60615573
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18462
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-60-615-573-18462

Query Match 4.8%; Score 169.4; DB 8; Length 180;

```
Best Local Similarity 96.6%; Pred. No. 1.3e-28; Mismatches 0; Indels 0; Gaps 0;
Matches 173; Conservative 0;

QY 2893 TCAGTTCTTAACAATAATGTAGTTCGCGCTTAATTCACCTAAATAATGGTACTAATGCTTCT 2952
D 2193 AGGAGCATGTTTTCATCTAGGAGCGTGGCCCATCTAGGAAACGTGGTTCATCTAGGAG 2252
QY 2611 TCGGATTTCAGATAGTATTCGCACTCCGACAGTCCGACAGTCCGATTCAGATAGGACTCAGAC 2670
D 2253 CATGCCCATCTAGGAGCGTGGCCCATCTAGGAAATGTGGTTCATCTAGGAGCGTGGC 2312
QY 2671 TCGGATAGCGACTCGGATTCAGATAGCGATTTCGGACTCAGATAGCGATTTCAGATTCAGAC 2730
D 2313 TCATCTAGGAGTGTGGTTCATCTAAGAGCGTGGCCCATCTAGGAGTGTGGTTCATCT 2372
QY 2731 AGCGATTTCAGAAATCAGACAGCGATTTCAGATTCA-----GACAGCGACTCAGACAGTGC 2784
D 2373 AGGAGCGTGGCCCATCTAGGAAATGTGGTTCATCTAGGAGTGTGGCTCATCTAGGAA 2432
QY 2785 TCAGATTTCAGATAGTACTCGGATTTCGCGAGTGTGATTCAGACTCAGTGTAGTACTCCGAT 2844
D 2433 TGTGGTTAATCTAGGAGCGTGGCCCATCTAGGAGCGTGGCCCATCTAGGAGTGTGGT 2492

RESULT 6
US-10-956-157-5134
; Sequence 5134, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5134
; LENGTH: 4050
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-5134

Query Match 4.0%; Score 141.4; DB 6; Length 4050;
Best Local Similarity 48.6%; Pred. No. 4.8e-22;
Matches 422; Conservative 0; Mismatches 441; Indels 6; Gaps 1;

QY 2011 AGCGATTCTAATTCAGATAGCGGTTTCAGATTCGGGTAGTGTATTCATCATCAGATAGTGT 2070
D 1653 AGGAGTGTGGTTCATCTAGGAGTGTGGCCCATCTAGGAAATGTGGTTCATCTAGGAG 1712
QY 2071 TCAGATTTCAGCGATGATTTCAGATTTCAGCAAGTGTTCAGACTCAGCGAGTGTTCAGAT 2130
D 1713 CGTGGCCCATCTAGGAGCGTGGTTCATCTAGGAGCGTGGCCCATCTAGGAGCGTGGC 1772
QY 2131 TCAGCAAGCGATTTCGACTCAGCGAGGATTCGCACTCAGCAATGACTTCGGATTTCAGAT 2190
D 1773 CCATCTAGGAGCGTGGTTCATCTAGGAGCGTGGCCCATCTAGGAGTGTGGTTCATCT 1832
QY 2191 AGCGATTTCAGTTCAGACAGTGTACTCAGATTTCGACAGTGTACTCAGATTTCAGATAGCGAT 2250
D 1833 AGGAGCGTGGCCCATCTAGGAAATGTGGTTCATCTAGGAGCGTGGCCCATCTAGGAG 1892
QY 2251 TCTGATCTCAGACAGTGTACTCAGATTTCAGATAGCGATTTCAGATTTCAGATAGCGATTTCAGAT 2310
D 1893 GGTGGTTCATCTAGGAGCGTGGTTCATCTAGGAGCGTGGCCCATCTAGGAGCGTGGC 1952
QY 2311 TCCGACAGTGTATTCGCACTCAGACAGCGATTTCGACTCCGACAGTGTATTCGCACTCAGAC 2370
D 1953 CCATCTAGGAGCGTGGCCCATCTAGGAGCGTGGCCCATCTAGGAGCGTGGCTCATCT 2012
QY 2371 AGCGATTTCAGTTCGACAGTGTATTCGACTCAGATAGCGATTTCGCACTCAGATAGCGAT 2430
D 2013 AGGAGCGTGGTTCATCTAGGAGCGTGGCCCATCTAGGAGCGTGGCCCATCTAGGAG 2072
QY 2431 TCAGATTTCAGACAGCGATTTCAGATTTCAGACAGCGATTTCAGATTTCAGATAGCGATTTCAGAT 2490
D 2073 CGTGGTTCATCTAGGAGCGTGGCCCATCTAGGAGCGTGGTTCATCTAGGAGTGTGGC 2132
QY 2491 TCCGACAGTGTACTCAGATTTCGCACTCAGATAGCGATTTCGCACTCAGATAGCGATTTCAGATTCGAC 2550
D 2491 TCCGACAGTGTACTCAGATTTCGCACTCAGATAGCGATTTCGCACTCAGATAGCGATTTCAGATTCGAC 2550
```

```
Db 2133 TCATCTAGGAAATGTGGTTAATCTAGGAGCATGGCCCATCTAGGAGCGTGGCTCATCT 2192
QY 2551 AGTCACTCAGATTTCGACAGTGTACTCAGACTCAGACAGTGTATTCGGATTTCAGCGATGAT 2610
D 2193 AGGAGCATGTTTTCATCTAGGAGCGTGGCCCATCTAGGAAACGTGGTTCATCTAGGAG 2252
QY 2611 TCGGATTTCAGATAGTATTCGCACTCCGACAGTCCGACAGTCCGATTCAGATAGGACTCAGAC 2670
D 2253 CATGCCCATCTAGGAGCGTGGCCCATCTAGGAAATGTGGTTCATCTAGGAGCGTGGC 2312
QY 2671 TCGGATAGCGACTCGGATTCAGATAGCGATTTCGGACTCAGATAGCGATTTCAGATTCAGAC 2730
D 2313 TCATCTAGGAGTGTGGTTCATCTAAGAGCGTGGCCCATCTAGGAGTGTGGTTCATCT 2372
QY 2731 AGCGATTTCAGAAATCAGACAGCGATTTCAGATTCA-----GACAGCGACTCAGACAGTGC 2784
D 2373 AGGAGCGTGGCCCATCTAGGAAATGTGGTTCATCTAGGAGTGTGGCTCATCTAGGAA 2432
QY 2785 TCAGATTTCAGATAGTACTCGGATTTCGCGAGTGTGATTCAGACTCAGTGTAGTACTCCGAT 2844
D 2433 TGTGGTTAATCTAGGAGCGTGGCCCATCTAGGAGCGTGGCCCATCTAGGAGTGTGGT 2492

RESULT 7
US-10-746-294A-30/c
; Sequence 30, Application US/10746294A
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)E

Query Match 2.8%; Score 99; DB 8; Length 171;
Best Local Similarity 81.9%; Pred. No. 3.2e-13;
Matches 127; Conservative 0; Mismatches 25; Indels 3; Gaps 1;

QY 2289 AGATTTCAGATAGCGATTTCAGATTTCGACAGTGTTCGCACTCAGACAGCGAT---TCTGA 2345
D 155 ATACTATGATGACGACTCAGATTTCAGATTTCAGATAGTGTGATTCAGACTCAGATAGCGACTCAGA 96
QY 2346 CTCGACAGTGTATTCGCACTCAGACAGCGATTTCAGATTTCGCACTCAGATAGTGTGATTCAGAT 2405
D 95 CTCGATAGCGATTTCGCACTCAGACAGCGACTCAGATTTCGCACTCAGATAGTGTGATTCAGATT 36
QY 2406 TAGCGATTTCGCACTCAGATAGCGACTCAGATTTCAG 2440
D 35 CAGTCACTCAGACTCAGATAGTGTGATTCAGATTTCAG 1

RESULT 7
US-10-746-294A-30/c
; Sequence 30, Application US/10746294A
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)E
```

```
; CURRENT APPLICATION NUMBER: US/10/746,294A
; CURRENT FILING DATE: 2003-12-29
; PRIOR APPLICATION NUMBER: US 09/692,412
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 09/803,736
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 124
; SEQ ID NO 30
; LENGTH: 98412
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-746-294A-30

Query Match      2.4%; Score 85; DB 6; Length 98412;
Best Local Similarity 48.0%; Pred. No. 3.2e-09;
Matches 275; Conservative 0; Mismatches 295; Indels 3; Gaps 1;

QY 2115 AGCGAGTGAATTCAGATTCAGCAAGCGGATTCGACTCAGCGAGCGATTCGCGACTCAGACAA 2174
   |||||
Db 51443 AGAGAGAGACTAGAGGCTGCAGGTGGCTGAGAGGGACAGGTTGATTGAAATGTCGGTGG 51384

QY 2175 TGACTCGGATTCAGATAGCGAATTCGACTCAGACAGTGAATTCAGATTCGCGACAGTGAATTC 2234
   |||||
Db 51383 TGGCTGAGACACTGGCAACGGCTTAGAGGTTGTAGGTGGCTGAGATACCGGCAATGGCTT 51324

QY 2235 AGATTGAGATAGCGATTCGACTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTC 2294
   |||||
Db 51323 AGAGGATTTGGGTGGCTGAGACACA---AATGGCTTAGAGGTCGTTGGTGGCTGAGAGAT 51267

QY 2295 AGATAGCGATTCAGATTCGAGACAGTGAATTCGAGACAGCGATTCGACTCCGACAG 2354
   |||||
Db 51266 AGGGAATGACTTAGAGGCGCTGGTGGCTGAGATGCTGGAAAGGCTTAGCGATCGATGA 51207

QY 2355 TGATTCGACTCAGACAGCGAATTCAGATTCGAGACAGTGAATTCGAGATTCAGATTCAGATTC 2414
   |||||
Db 51206 TGGTTGAGACATCGAGAAATGGCTGAGAAATTTGGTCTTGGTTGGGACACCGGGAACGGTT 51147

QY 2415 CGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGATTC 2474
   |||||
Db 51146 AGAAGCGGGTTGGCTGAGACACCGGAAATGGCTTAGAGTTTGGATGGCTGAGACAC 51087

QY 2475 AGATAGCGATTCAGATTCGAGACAGTGAATTCGAGACAGTGAATTCGAGATTCAGATTCAGATTC 2534
   |||||
Db 51086 CGGGAATGGCTTAGAGTTTGTATGACTGAGACCGGCAATGGCTTTGACCTTTGTTGA 51027

QY 2535 CGATTGAGATTCGAGACAGTGAATTCGAGACAGTGAATTCGAGATTCAGATTCAGATTCAGATTC 2594
   |||||
Db 51026 TGACTGAGAAACCGACAAAGGCTTAGACITTTTGTGATGACTGAGACACTGGGAATGGCTT 50967

QY 2595 GGATTCAGCGAGTGAATTCGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2654
   |||||
Db 50966 AGATTTTGGATGACTGAGACACTCGGAACGGCTTAGAGTTTCGATGGCTGAGACAA 50907

QY 2655 AGATAGCGACTCAGACTCGGATAGCGACTCGGA 2687
   |||||
Db 50906 TGGAAATGGTTAGAGGTCGGTGGCTGAGA 50874

RESULT 8
US-60-615-573-18342
; Sequence 18342, Application US/60615573
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18342
; LENGTH: 2701
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-60-615-573-18342
```

```
Query Match      2.3%; Score 79.8; DB 8; Length 2701;
Best Local Similarity 43.3%; Pred. No. 1.3e-08;
Matches 393; Conservative 25; Mismatches 477; Indels 12; Gaps 3;

QY 606 ATCAACAAATGCAATGCAAGGAAACCGCGGTAACTGGTGAAGCTACTACTACGACAAC 665
   |||||
Db 468 AGCACCATAATWATACAAATAACGATGTAACACCACTCTACAAGTGAACCATCTAC 527

QY 666 GAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGCGAGGAATPAGT 725
   |||||
Db 528 AAGTGAATTTCAAAACAAACCAACTACCTCAAGAATCTTACAAATATTGAAATTTCA 587

QY 726 GAATCAACCAAGTAATGAACGACTTTTATGATATACTAATACAGATATCATCTGTAATTC 785
   |||||
Db 588 ACCGCAACCAACCGCTTCAAAAGTAGACAATCAAGTTACAGATGCAACTAATCCAAAGA 647

QY 786 ACCTCAAAATTTCAAAATCGGGAATAATGTTTCAACAAACGCAAGATACTTCAACTGAAGC 845
   |||||
Db 648 ACCAGTAATGTCTCAAAAGAGAACTTAAANAAATCTTGAGAAATTAAGANTRGT 707

QY 846 AACACCTTCAACAAATGAATCAGCTCCACAGATGACAGATGACAGTAATGAAGATPAGT 905
   |||||
Db 708 TAGAAATGATFARCAATACAGATCTTTCACTAAACCAAGTTGCTTACAGCTCCAAACAGT 767

QY 906 TAATCAAGCGGTAAATACAGTGGCTAGAAATGAGAGCAATTTAGTTTAGCGGCAAGTAC 965
   |||||
Db 768 TGCACCAAAACGKTAATGVCRAAATGCGYTTTGCAAGTTGCAACACGACGCGTTC 827

QY 966 TGCAGATGCAACCGGCAAGCTGGCAAGATTTAGCAATCAGTTACGCAATGTGACAGTTGG 1025
   |||||
Db 828 TTCAAAATGTAATGATTAATTAATTAAGTACGAAACACACATCAAAAGTTGGCGATGG 887

QY 1026 TATGACTCTGGTACGACTGTGTATCCGACCAAGCAGGTTATGTCAAACTGAATATGG 1085
   |||||
Db 888 TAAAGATAATGTGGCAGCAG-----CGCATGACCGTAAAGATATTGAATATGATACAGA 941

QY 1086 TTTTTCAGTGCCTTAATCTGCTGTTAAAGGTGACACATTCAAATTAACCTGACTACCTAAGA 1145
   |||||
Db 942 GTTTACAATTGACATAAAGTCAAAAGGCGGATCAATGACGATTAATTAATGATAAGA 1001

QY 1146 ATTAAACTTTAAATGGTGTAAAC---TTCAACTGCTAAAGTGGCCACCAATTTATGGCTGAGA 1202
   |||||
Db 1002 TGTAAATTCCTCGGATTTAAACAGATAAAATGATCCTATCGATATTACTGATCCATCAG 1061

QY 1203 TCAGTATTGGCAATATGGTGTATTCGATAGTGAATGTAAGT---TTATTTATACATTTAC 1259
   |||||
Db 1062 AGAGGTTCATTGCTGAAAGGAAACATTTGATTAAGGCACTAAGCAAAATCACATATACATTTAC 1121

QY 1260 AGACTATGTAATTAATCAAGATGATGTAAGCAACTTTGACCACTCCCGCTTTATTTGA 1319
   |||||
Db 1122 AGATTAATGATTAATTAATGAAGATATAAAACACGCTTAACTTATATCTATATTGA 1181

QY 1320 CCTGAAATGTTAAAGACAGAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAAAC 1379
   |||||
Db 1182 TAARMAARCAGTWCWAAATGACACWAGTTTGAATTTAACTTTGACAGCAGGTAAGA 1241

QY 1380 AGCAACAAACACAGTATTAGTATGATAAGAAAAATATGTAAGTTTATTAACATTTATCTAT 1439
   |||||
Db 1242 AACWAGCCAAAGTGTGTTGATTTATCAAGAYCCAATGGTYCATGTTGATTTCAACAT 1301

QY 1440 TAAAGGTACAATTTGACCAATCGATAAACAATTAATAGTATCGTACAGCAATTTATGTT 1499
   |||||
Db 1302 TCAATCTATCTTTACAAATTTAGATGAARAYAACAACTATTGAACACAAATTTATGTT 1361

QY 1500 CAATCCA 1506
   |||||
```


Db 1362 TAAVCCW 1368

RESULT 9

US-10-220-366A-1506
; Sequence 1506, Application US/10220366A

; GENERAL INFORMATION:

; APPLICANT: HYSEQ, INC

; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

; FILE REFERENCE: 21272-042

; CURRENT APPLICATION NUMBER: US/10/220,366A

; CURRENT FILING DATE: 2002-08-28

; PRIOR APPLICATION NUMBER: 09/577,409

; PRIOR FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: 09/515,126

; PRIOR FILING DATE: 2001-02-23

; NUMBER OF SEQ ID NOS: 27802

; SOFTWARE: Custom

; SEQ ID NO 1506

; LENGTH: 388

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-220-366A-1506

Query Match 2.3%; Score 79.4; DB 6; Length 388;

Best Local Similarity 50.7%; Pred. No. 8.2e-09;

Matches 191; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 2175 TGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTCACTCAGATTCGGACAGTCACTC 2234

Db 9 TGGGACGGAGAGAGATACAGTTTATAGAGAGAGACGGAGAGAGAGACAGAGGGAGAGAG 68

QY 2235 AGATTCAGATAGCGATTCTGACTCAGACAGTCACTCAGATTCAGATTCAGATTC 2294

Db 69 GGATACAGTGGAGATACAGGAGAGAGAGACATCGGAGAGAGAGATAGGTTTAG 128

QY 2295 AGATAGCGATTCAAGATTCGGACAGTCACTCGGACTCAGACAGCGATTCTGACTCCGACAG 2354

Db 129 AGAGACAGATGAGATGGAGACAGAGTATAGACAGGGAGAGAGACAGAGTGG 188

QY 2355 TGATTCGACTCAGACAGCGATTCACTCAGATTCGGACAGTCACTCAGATTCGAGTTC 2414

Db 189 AGAGAGAGACTGAGGCTGAGAGAGACTCGGAGAGAGATCTTATATTGTGAGGGAGAGAG 248

QY 2415 CGACTCAGATCGGACTCAGATTCAGACAGCGATTCACTCAGATTCAGATTCAGATTC 2474

Db 249 AGACTGAGAGAGAGACAGAGGAGAGAGCGATATAGAGAGGGAGAGAGATTTGTTAGAG 308

QY 2475 AGATAGCGATTCAAGATTCGGACAGTCACTCAGATTCGGACAGTCACTCAGATTCAGATAG 2534

Db 309 ATAGATATATGGAGATAGATAGAGACATTTTGAGAGAGTGAAGAGGGATAGAGAGAC 368

QY 2535 CGATTCAGATTCGGACA 2551

Db 369 AGCGTTAGATTCTGAGA 385

RESULT 10

US-10-220-366A-1495

; Sequence 1495, Application US/10220366A

; GENERAL INFORMATION:

; APPLICANT: HYSEQ, INC

; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

; FILE REFERENCE: 21272-042

; CURRENT APPLICATION NUMBER: US/10/220,366A

; CURRENT FILING DATE: 2002-08-28

; PRIOR APPLICATION NUMBER: 09/577,409

; PRIOR FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: 09/515,126

; PRIOR FILING DATE: 2001-02-23

; NUMBER OF SEQ ID NOS: 27802

; SOFTWARE: Custom

; SEQ ID NO 1495

; LENGTH: 427

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-220-366A-1495

Query Match 2.1%; Score 73.8; DB 6; Length 427;

Best Local Similarity 53.3%; Pred. No. 1.4e-07;

Matches 203; Conservative 0; Mismatches 172; Indels 6; Gaps 2;

QY 2279 ATAGCGATTCAAGATTCAGATAGCGATTCCGACAGTCACTCAGATTCGGACAGTCACTCAGACGCG 2338

Db 47 AGAGGGAGACAGACAGACAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 106

QY 2339 ATTCTGACTCCGACAGTCACTCAGATTCGGACAGTCACTCAGATTCGGACAGTCACTCAG 2398

Db 107 GCAGGTGACGACAGAGGGCGGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 166

QY 2399 ACTCAGATAGCGATTCCGACTCAGATAGCGACTCAGATTCAGATTCAGACAGCGATTCACTCAG 2458

Db 167 ACAGATAGACAGGCTGAGGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 222

QY 2459 ACAGCGATTCAAGATTCAGATAGCGATTCCGACAGTCACTCAGATTCGGACAGTCACTCAG 2518

Db 223 AGACAGACAG 282

QY 2519 ACTCGGATTCAAGATTCAGATTCGGACAGTCACTCAGATTCGGACAGTCACTCAGATTCGAC 2576

Db 283 ACAGAGGTAG 342

QY 2577 AGACTCAGACAGTCACTCAGATTCGGACAGTCACTCAGATTCGGACAGTCACTCAGATTCG 2636

Db 343 AGACAGAGACAG 402

QY 2637 CGACAGTCACTCGGATTCCAGA 2657

Db 403 TGAGACATATCTTAGACAGA 423

RESULT 11

US-60-615-573-14147

; Sequence 14147, Application US/60615573

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Mounts, William M

; APPLICANT: Murphy, Ellen

; APPLICANT: Olmsted, Stephen

; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES

; FILE REFERENCE: 031896-084099 (AM 101724)

; CURRENT APPLICATION NUMBER: US/60/615,573

; CURRENT FILING DATE: 2004-10-05

; NUMBER OF SEQ ID NOS: 18598

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 14147

; LENGTH: 2493

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-60-615-573-14147

Query Match

Best Local Similarity 49.1%; Pred. No. 4.7e-07;

Matches 252; Conservative 10; Mismatches 239; Indels 12; Gaps 3;

QY 1244 TTATTTATACATTTTACAGACTATGTAATAATCTAAAGATGATGTAAGCACTTTGACCA 1303

Db 1112 TYACCTATACITTTTACAGATTATGATATAATGATAATTAAGAGCACTTTAAAT 1171

QY 1304 TGCCCGCTTATTTGACCTGAAAATGTT---AAAAAGACAGTGAATGTGACATTTGGCTA 1360

Db 1172 TAACTCATCATTTGATAAATCAAAGGTTCCAAATATAAAYACAAAGTTAGATGAGAT 1231

QY 1361 CTGGCATAGGTAGTACACAGCAAAACAAAGTATTAGTATGATTAATAAATATGGTA 1420

Db 1232 ATAARACGGCCCTTTTCATCATGTAATTAACAAATACGCTTGAATATCAWRACCTAACG 1291

1421 AGTTTATACTTATCTATTAAAGGTACAAATGACCAAAATCGATAAAACAAATAATACGT 1480
1292 AAAATCGGACTGTAACCTTCAAGATGATGTTTACAAATATAGATACGAAATTAATACAG 1351
1481 ATCGTCAGACAAATTTATGTCATCAATCCAAAGTGGAGATAACGTTATTGCGCGGTTTAAACAG 1540
1352 TTGAGCAAAAGGATTTATATTAATTAACCTCTCTGTTATTCAGCAAGAAACAAATGTA---- 1407
1541 GTAATTAAACCAATACCGATAGTAAATGCAATTAATAGATACCAAAATACAGATTA 1600
1408 --AATATTTTTCAGGAATGGGATGAAGGTTCAACAATTAATGACATAGTACCAATTA 1465
1601 AAGTATATAAGTAGATAAATGCACTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAA 1660
1466 AAGTTTATAAGTTGGAGATAATCAAAATTTACAGATAGTACAGAAATTTATGATTA 1525
1661 ---ACTTTGAGGATGTCACATAATAGTGTGAATATTAATTCATTCACAAATCCAAATCAATATA 1717
1526 GTCAATATGAAGATGTCACAAATGATGATTATGCCAAATTAGGAAATTAATGAYTGA 1585
1718 AAGTAGATTTTAACCCCTGATGATCAAAATTA 1750
1586 ATATTAAATTTTGGTAAATATAGATTCACCAATATA 1618

RESULT 12

US-60-613-292-360

; Sequence 360, Application US/60613292
; GENERAL INFORMATION:
; APPLICANT: Gene Logic, Inc
; APPLICANT: HIGGS, Brandon
; APPLICANT: ELASHOFF, Michael
; APPLICANT: MENDRICK, Donna L.
; APPLICANT: PORTER, Mark W.
; APPLICANT: CASTLE, Arthur L.
; APPLICANT: JOHNSON, Kory R.
; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
; CURRENT APPLICATION NUMBER: US/60/613,292
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 1300
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 360
; LENGTH: 5192
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION:
US-60-613-292-360

Query Match 2.0%; Score 71.6; DB 8; Length 5192;
Best Local Similarity 52.7%; Pred. No. 1e-06;
Matches 155; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
2429 ACTCAGATTCAGACAGCGATTCAGATTCACACAGCGATTCAGATTCAGATAGCGATTCAG 2488
3785 ACACAGAGACAGACAGACAGACAGACAGACAGACAGACAGATAGACAGAGATAGAC 3844
2489 ATTCGACAGTACTCAGATTCGACAGTGAATCGGATTCAGATAGCGATTCAGATTCG 2548
3845 AGATAGATGACAGAGAGAGCGGACAGAGACAGAGACAGAGAGAGACACAGAGAAAGAG 3904
2549 ACAGTACTCAGATTCGACAGTGAATTCAGATTCAGACAGTGAATTCGATTCAGCGAGTG 2608
3905 AGAGAGAGAGAGAGAGAGACAGACAGACAGACAGACAGACAGAGAGAGAGAGAGAGAG 3964
2609 ATTCGATTCAGATAGTGAATTCGACTCCGACAGTGAATTCGATTCAGATAGCGACTCAG 2668
3965 ATAGACAGACAGAGACAG 4024
2669 ACTCGGATAGCGACTCGGATTCAGATAGCGATTCGAGTTCGATAGCGATTCAG 2722

Db 4025 AGAGGAGAGGGAGAGAGGGAGAGAGAGAGAGAGAGAGGGAGGGAGAGAGAGAG 4078

RESULT 13

US-60-615-573-14144
; Sequence 14144, Application US/60615573
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M.
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14144
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-60-615-573-14144

Query Match 2.0%; Score 71.2; DB 8; Length 404;
Best Local Similarity 51.6%; Pred. No. 5.2e-07;
Matches 160; Conservative 1; Mismatches 149; Indels 0; Gaps 0;
1985 ATTCTGACCAGGTTTCAGATTCGGCAGCGATTCCTAATTCAGATAGCGGTTTCAGATTCGG 2044
1 AAYTTGTACCAGAGTCTGAATCACTATCTGAATCTGARTCACTCTGARTCTGAAATCAC 60
2045 GTAGTGATTCATACATAGATGTTGTTTCAGATTCAGGAGTTCAGATTCAGATTCAGATTCAG 2104
61 TATCTGAATCTGAATCACTCTGAGTCTGARTCACTCTGATCTGAAATCTGARTCACTATCTG 120
2105 ATTCTGAGTCTGAGTCTGATTCAGATTCAGATTCAGCAAGCGATTCGAGTCTGAGGAGGATTCGG 2164
121 ARTCTGAATCACTGTTCTGAATCTGAGTCACTATCTGAATCTGAGTCACTATCTGAAATCTG 180
2165 ACTCAGACAAATGACTCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2224
181 AGTCACTGTTCTGAATCTGAATCACTATCTGAATCTGAATCACTATCTGAAATCTGAAATCAC 240
2225 ACAGTCACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2284
241 TATCTGAGTCTGAATCACTATCTGAATATATGTCATTTCTGAGTCACTATTTTGAACACGA 300
2285 ATTCTGATTC 2294
301 TATCTGATTC 310

RESULT 14

US-10-746-294A-111
; Sequence 111, Application US/10746294A
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)E
; CURRENT APPLICATION NUMBER: US/10/746,294A
; CURRENT FILING DATE: 2003-12-29
; PRIOR APPLICATION NUMBER: US 09/692,412
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 09/803,736
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 124
; SEQ ID NO 111
; LENGTH: 105306
; TYPE: DNA

; ORGANISM: Arabidopsis thaliana
US-10-746-294A-111

Query Match 1.9%; Score 67.2; DB 6; Length 105306;
Best Local Similarity 48.9%; Pred. No. 2.5e-05;
Matches 180; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

```
QY 2233 TCAGATTTCAGATGCGGATTCCTGACTCAGACAGTGAATCCGAGTTCAGATGCGGATTCAGAT 2292
DB 63644 TCAGATTCCTTTTCAGATTCCTCTCCAAATCATTCCTCAGAAATCCTTTTTCAGATTCCTCT 63703
QY 2293 TCAGATGCGGATTCAGATTCGACAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2352
DB 63704 TCAGATTCATATTCATATTCCTCTCTAGATCCCTCTCCGATCCCTTTTCAGATTCCTCT 63763
QY 2353 AGTGATTCGAGTTCAGACAGGATTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2412
DB 63764 TCAGATTCATATTCAGAACTCTCTCTAGGTCCTCTCTAGGTCCTCTCTAGGTCCTCTCTAG 63823
QY 2413 TCCGACTCAGATGCGGATTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2472
DB 63824 TCAGATTCATATTCAGAACTCTCTCTAGGTCCTCTCTAGGTCCTCTCTAGGTCCTCTCTAG 63883
QY 2473 TCAGATGCGGATTCAGATTCGACAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 2532
DB 63884 TCAGATTCATATTCAGAACTCTCTCTAGGTCCTCTCTAGGTCCTCTCTAGGTCCTCTCTAG 63943
QY 2533 AGCGATTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCG 2592
DB 63944 TTATATCCCTCTTCAGATTCCTTTTATATACCTCTTCAGATTCATATTCATATCCCTCT 64003
QY 2593 TCGGATTC 2600
DB 64004 TTAGATCC 64011
```

RESULT 15

US-60-615-573-18593
; Sequence 18593; Application US/60615573

GENERAL INFORMATION:

; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18593
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-60-615-573-18593

Query Match 1.9%; Score 67; DB 8; Length 2022;
Best Local Similarity 37.8%; Pred. No. 7.4e-06;
Matches 372; Conservative 93; Mismatches 499; Indels 21; Gaps 4;

```
QY 541 AACACTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCACAAACGAGGAAACGACACAA 600
DB 997 AAAACWGCATYGAAGCAATGAATTCGACTGTGAAGAACACAAAGCAGCGAAAGAH 1056
QY 601 TCATCATCAAAATGCACTACGGAAGAAACGCGGTA---ACTGGTGAAGCTACTACT 657
DB 1057 AAAATGKATCAAGCGWTGKMTGCAACCGGTGATTCGATTCGATTCGATTCGATTCGATTCGAT 1116
QY 658 ACGACAAACGAATCAAGCTAATCAACCGCAACCACTCAATCAAGCAATCAAAATCGGGAG 717
DB 1117 GATGTGTGATGCAAAACTCAAAATGAAGTACAAATCGGAGCCATACACCTGATGCA 1176
QY 718 GAATTAGTGAATCAAAACAGTAATGAACGACTTTTAAATGATTAATCAATCAATCAATCAATCT 777
```

```
DB 1177 AATGTTAAACWWSMAGCRRAAACAGCAATTCAGATAAAGTACAAGCKCAAGAAAMAGCA 1236
QY 778 GTAATTCACCTCAAAATTCCTACAAATCGCGAAATGTTTCAACAAACGCAAGATCTTCA 837
DB 1237 ATTGATGWSAATAACGGYKCAACACACGAAACAAACAGCAGCAGCAACAAACAGTCA 1296
QY 838 ACTGAAGCAACACCTTCAACAAATGAATTCAGCTCCACAGAGTACAGATGCAAGTAATAA 897
DB 1297 ACWGAACAAACACACGCTGATTCAGCAATAGATGWSGCACATWCAATGCR----- 1347
QY 898 GATGTAGTTAATCAAGCGGTTTAATCAAGTCCGCTAGAAATGAGAGCAATTTAGTTAGCG 957
DB 1348 ---GAAGTTGAAGCGGCTTAARAACGCAAAATTCCTAAATTTGAAGCTATTCARCCAGC 1404
QY 958 GCAGTAGCTGAGATGACACCGGAGCTGGCAGACAGATTTACGAATCAGATTCAGCAATG 1017
DB 1405 ACAACAAACAAAGATATTCGCAAAACAGCAATTCAGCAAGCAATTCAGCAAGCAATTC 1464
QY 1018 ACAGTTGTTATTTGACTCTGCTGACGACTGTATCCGACCAAGCAGGTTTATGTCAAACTG 1077
DB 1465 GCMATYGTCAACCGCAGASATTCACWGCWGAAGAAATTCGACGGCTAATGCGRACGTA 1524
QY 1078 AATTATGTTTTCAGTCCCTTAATCTCTGTTTAAAGTGACACATTCACAAATTAATCTA 1137
DB 1525 GATAATGCTGTRACASAAAGCAAAAYARCMACATTCGARRCTGCTAATAGTCAAAATGAWGT 1584
QY 1138 CCTAAGAAATTAACCTTAATGTTGTTAACTTCACTGCTAAAGTCCCAACCAATTTATGCT 1197
DB 1585 GAYCAAG---CGAARACACACAGGTGAATAGTATTTGATCAAGTRACACCAACGTTTAT 1641
QY 1198 GGAGATCAAGTATTGGCAAAATGGTGTATTAATCGATGATGTTGTTATTTATACATTT 1257
DB 1642 AAAAAGCWAACGACGTAATGAATCACWCAATTTTAAATAACAARTTRCAAGMRAAT 1701
QY 1258 ACAGACTATTAATACTAAAGATGATGTAAGCAACTTTTGACCATGCGCCCTTATAT 1317
DB 1702 CAAGCTACGCCAGATGCAACAGATGAGARAAAACAGCAGCAGWAGWRCWAGCAAGTACT 1761
QY 1318 GACCTTGAAATGTTTAAAGACAGGTAATGTGACATTTGGCTTACTGGCATAGGTAGTACA 1377
DB 1762 GAAATGTTAAAGCAATTCAGCCATYRCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1821
QY 1378 ACAGCAAAACAAACAGTATTAGTATTGATAAATAATGTTAAGTTT---TATACTTA 1434
DB 1822 GCWAAAACWAAATGCAGAGACGCRATTAAAGCKGTACACCAACAAAGTTTGTGAAGAAACAA 1881
QY 1435 TCTATTAAGGTACAATTCACCAATCGATTAACAAACAAATTAATACGTATCGTCAGCAAT 1494
DB 1882 RCRGCKAAAGATGARATYATTCATTAACAGCARCRAACARMKGTATTATTAATTAAYGAT 1941
QY 1495 TATGTCAATCCAAAGTGAGATAACG 1519
DB 1942 CARAAGCAGCAATGTAAGAAAG 1966
```

Search completed: November 5, 2004, 16:28:26
Job time : 122 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2004, 05:31:43 ; Search time 10221 Seconds
(without alignments)
12471.009 Million cell updates/sec

Title: US-09-679-643-1

Perfect score: 3498

Sequence: 1 ggtaacataataacacatc.....gaagagtataagaagaagctt 3498

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:.*
2: gb_est2:.*
3: gb_hic:.*
4: gb_est3:.*
5: gb_est4:.*
6: gb_est5:.*
7: gb_est6:.*
8: gb_gss1:.*
9: gb_gss2:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	740.8	21.2	768	9	CC958014 BOIHQ26TR
C 2	518.2	14.8	534	9	CC942933 BOIHQ26TF
C 3	212.4	6.1	697	5	BW250931 BW250931
C 4	205.2	5.9	772	8	CC099392 CSU-K34.1
C 5	205	5.9	773	8	CC099378 CSU-K34.1
C 6	182.4	5.2	811	7	CO426942 UI-M-HUO
C 7	179.8	5.1	667	5	BW254640 BW254640
C 8	171.8	4.9	699	5	BX729701 BX729701
C 9	167.6	4.8	928	9	CL129761 ISB1-97D4
C 10	167.4	4.8	563	7	CN689123 E0270H03-
C 11	164.6	4.7	817	9	CC507028 CH240_348
C 12	163	4.7	625	1	AV992680 AV992680
C 13	162.8	4.7	544	7	CN688762 E0264F04-
C 14	162.2	4.6	973	7	CK423535 AUF Ipspn
C 15	160.2	4.6	1526	1	AL162088 DRFZp7611
C 16	158.6	4.5	607	7	CF894873 A0141B10-
C 17	158.4	4.5	737	8	BZ191186 CH230-424
C 18	157.4	4.5	870	5	BX078600 BX078600
C 19	156.8	4.5	542	8	AZ875941 2M0190D14
C 20	156.6	4.5	623	8	CC075747 CSU-K33r.
C 21	155.2	4.4	542	8	AZ875941 2M0190D14
C 22	152.4	4.4	941	7	CK425697 AUF Ipsps
C 23	150.6	4.3	599	6	CB582745 AMGNNUC:N
C 24	149	4.3	652	5	BW388427 BW388427

25	148.8	4.3	944	9	CNS04SSE	AL305591 Tetraodon
C 26	148.6	4.2	573	7	CF893788	CF893788 A0127F06-
C 27	146.8	4.2	744	5	EX078599	EX078599 EX078599
C 28	146.6	4.2	860	9	CNS02E27	AL193192 Tetraodon
C 29	146	4.2	564	1	AV854983	AV854983 AV854983
C 30	145.2	4.2	927	7	CK420227	CK420227 AUF Ipsps
C 31	145	4.1	704	7	CF899579	CF899579 A0304H06-
C 32	141.2	4.0	680	7	CK344793	CK344793 K0975C05-
C 33	141.2	4.0	984	9	CNS071N7	AL425141 clone BAO
C 34	140.6	4.0	650	9	AG273016	AG273016 Cyanidios
C 35	140.4	4.0	623	8	CC075747	CC075747 CSU-K33r.
C 36	140	4.0	950	7	CK425370	CK425370 AUF Ipsps
C 37	139.4	4.0	832	4	BM170146	BM170146 EST572669
C 38	139.2	4.0	723	9	CNS01U74	AL168241 Tetraodon
C 39	137.8	3.9	1939	3	AK079258	AK079258 Mus muscu
C 40	136.8	3.9	571	8	BZ214163	BZ214163 CH230-331
C 41	136.2	3.9	773	6	CB939384	CB939384 IPCGX14
C 42	136	3.9	764	2	BF024332	BF024332 Pap 058 L
C 43	135.4	3.9	960	9	CNS01EB6	AL140431 Anopheles
C 44	135.2	3.9	891	5	BM120795	BM120795
C 45	134.4	3.8	554	6	CB613746	CB613746 AMGNNUC:N

ALIGNMENTS

RESULT 1
CC958014/c
LOCUS BOIHQ26TR BO 1.4 1.6 KB nuc Brassica oleracea genomic clone linear GSS 18-AUG-2003
DEFINITION BOIHQ26, genomic survey sequence.
ACCESSION CC958014
VERSION CC958014.1 GI:33801933
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE
AUTHORS Town, C.D., Van Aken, S., Usterback, T., Koc, H. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOIHQ26TF
Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
source 1. .768
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone_lib="BO 1.4 1.6 KB nuc"
/note="Vector: pHOS2; Site 1: BstXI; 1.4-1.6 kb sheared nuclear DNA inserted into pHOS2 using BstXI linkers"

FEATURES
Query Match 21.2%; Score 740.8; DB 9; Length 768;
Best Local Similarity 97.8%; Pred. No. 1.1e-138;
Matches 751; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
ORIGIN
1275 TAAAGATGATGTAAGACCACTTTGACCGCTTATATGACCGCTTAAATCTTAA 1334
768 TAAAGATGATGTTAAAGACCTTTTGACCGCTTATATGACCGCTTAAATCTTAC 709

1335 AAAGACAGTAAATGTGACATTCGGCTACTGCGATAGGTAGTCAACAGCAAAACAAAACAGT 1394
 1395 ATTAGTAGATATGAAAAATATGTAAGTATTTATTAACCTTATCTATTAAGGTGCAATTTGA 1454
 648 ATTAGTAGATATGAAAAATATGTAAGTATTTATTAACCTTATCTATTAAGGTGCAATTTGA 589
 1455 CCAATCGATTAACAAATATATAGTATCGTCGACAGCAATTTATGTCATCAACCAAGTGGAGA 1514
 588 CCAATCGATTAACAAATATATAGTATCGTCGACAGCAATTTATGTCATCAACCAAGTGGAGA 529
 1515 TAAAGTTATGCGCGGTTTAAACAGGTAAATTTAAACCAATATCGGATAGTAAAGCAATT 1574
 528 TAAAGTTATGCGCGGTTTAAACAGGTAAATTTAAACCAATATCGGATAGTAAAGCAATT 469
 1575 AATAGATCAGCAAAATACAGGTATTTAAAGTATATTAAGTATAGTATGAGTATTTATTC 1634
 468 AATAGATCAGCAAAATACAGGTATTTAAAGTATATTAAGTATAGTATGAGTATTTATTC 409
 1635 TGAAGTTATGCGCGGTTTAAACAGGTAAATTTAAACCAATATCGGATAGTAAAGCAATT 1694
 408 TGAAGTTATGCGCGGTTTAAACAGGTAAATTTAAACCAATATCGGATAGTAAAGCAATT 349
 1695 ATTCCCAATCAATCAATATTAAGTATAGTATTTAAACCGCTGATGATCAAAATTTACAAC 1754
 348 ATTCCCAATCAATCAATATTAAGTATAGTATTTAAACCGCTGATGATCAAAATTTACAAC 289
 1755 ACCGTATATAGTATGTTTAAAGTATGATCGGATAGTATTTAAACCGCTGATGATCAAAATTTACAAC 1814
 288 ACCGTATATAGTATGTTTAAAGTATGATCGGATAGTATTTAAACCGCTGATGATCAAAATTTACAAC 229
 1815 ACCTTCACCTTATATGCGGTATTAACCGATATATTTGGCGCTCTATGTCATGGGACAA 1874
 228 ACCTTCACCTTATATGCGGTATTAACCGATATATTTGGCGCTCTATGTCATGGGACAA 169
 1875 CGAAGTACATTTAATAACCGGTATGAGTTCTGGTACCGGTATCGGATTAACCGGTATGTTCC 1934
 168 CGAAGTACATTTAATAACCGGTATGAGTTCTGGTACCGGTATCGGATTAACCGGTATGTTCC 109
 1935 TGAACAACTGATGAGCTGCTGAAATTTGAACCAATTCAGAGGATTCAGATTTCTGACCC 1994
 108 TGAACAACTGATGAGCTGCTGAAATTTGAACCAATTCAGAGGATTCAGATTTCTGACCC 49
 1995 AGGTTACAGTTCTGCGAGGTTCTTAATTCAGATGCGGTTTCAAGATTC 2042
 48 AGGTTACAGTTCTGCGAGGTTCTTAATTCAGATGCGGTTTCAAGATTC 1

CC942933 534 bp DNA linear GSS 18-AUG-2003
 LOCUS BOHQ26TF BO_1.4_1.6 KB nuc Brassica oleracea genomic clone
 DEFINITION BOHQ26, genomic survey sequence.
 CC942933
 ACCESSION CC942933.1 GI:33775799
 VERSION GSS.
 KEYWORDS Brassica oleracea
 SOURCE Brassica oleracea
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 534)
 Town, C.D., V. Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished (2001)
 Other GSSs: BOHQ26TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: Sheared ends.
 Location/Qualifiers
 1. 534
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOHQ26"
 /clone_lib="BO_1.4_1.6 KB nuc"
 /note="Vector: pHOS2; Site 1: BstXI; 1.4-1.6 kb sheared nuclear DNA inserted into pHOS2 using BstXI linkers"

ORIGIN

Query Match 14.8%; Score 518.2; DB 9; Length 534;
 Best Local Similarity 99.3%; Pred. No. 7.4e-94;
 Matches 531; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 418 AGTGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATTTCA 477
 Db 1 AGTGAATAATAGT-TTCGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATTTCA 59
 QY 478 AGTAGCGTTAGTGTGACCTTAAACAGACACAAACGTTAGTGTATCTAAACATCG 537
 Db 60 AGTAGCGTTAGTGTGACCTTAAACAGACACAAACGTTAGTGTATCTAAACATCG 119
 QY 538 TCAACACTAATAATGCGGAACGAGTGTGCGCAAAATCCAGCACACACGAAACGACA 597
 Db 120 TCAACACTAATAATGCGGAACGAGTGTGCGCAAAATCCAGCACACACGAAACGACA 179
 QY 598 CAATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAATCTGGTGAAGTACTACT 657
 Db 180 CAATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAATCTGGTGAAGTACTACT 239
 QY 658 ACGCAACGAAATCAAGTAAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAG 717
 Db 240 ACGCAACGAAATCAAGTAAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAG 299
 QY 718 GAATAGTAGAATCAACAAAGTAATGAACGACTTTTATGATCTACTTAATACAGTATCACT 777
 Db 300 GAATAGTAGAATCAACAAAGTAATGAACGACTTTTATGATCTACTTAATACAGTATCACT 359
 QY 778 GTAAATTCACCTCAAAATTTCTCAAAATCGGAAATGTTTCAACAAACGAAATACTTCA 837
 Db 360 GTAAATTCACCTCAAAATTTCTCAAAATCGGAAATGTTTCAACAAACGAAATACTTCA 419
 QY 838 ACTGAAGCAACACCTTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAAATA 897
 Db 420 ACTGAAGCAACACCTTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAAATA 479
 QY 898 GATGTAGTAAATCAGCGGTTAATACAGTGGCCCTAGATGAGAGCAATTTAGTT 952
 Db 480 GATGTAGTAAATCAGCGGTTAATACAGTGGCCCTAGATGAGAGCAATTTAGTT 534

RESULT 3
 BW250931 697 bp mRNA linear EST 09-NOV-2002
 LOCUS BW250931 Nori Satoh unpublished cDNA library, tailbud embryo Clona
 DEFINITION intestinalis cDNA clone citb08h22 5', mRNA sequence.
 CC942933
 ACCESSION BW250931.1 GI:24830849
 VERSION EST.
 KEYWORDS Ciona intestinalis
 SOURCE Ciona intestinalis
 ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.
 1 (bases 1 to 697)
 Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N.
 Expressed genes in Ciona intestinalis (2002c)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Nori Satoh


```

Db      348 GATTGGGATTTGGGATTTGGGATTTGGGATTTGGGATTTGGGATTTGGGATTTGG 289

Qy      2656 GATAGCAGCTCAGACTCGGTAGCGACTCCGATTTCAGATACGATTCGGACTCAGATPAGC 2715
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      288 GATTGGGATTTGGGATTTGGGATTTGGGATTTGGGATTTGGGATTTGGGATTTGG 229

Qy      2716 GATTCAGAATCAGA 2729
          ||||| ||||| |||||

Db      228 GATTGGGATTTGGGA 215
          ||||| ||||| |||||

RESULT 5
CC099378/c
LOCUS    CC099378              773 bp     DNA         linear   GSS 16-APR-2003
DEFINITION
CSU-K34.113M15.SP6 CSU-K34 Aedes aegypti genomic clone
CSU-K34.113M15, genomic survey sequence.
ACCESSION
CC099378
VERSION  CC099378.1 GI:29965941
KEYWORDS GSS.
SOURCE   Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;
          Stegomyia.
REFERENCE
1 (bases 1 to 773)
Loftus,B., Shetty,J., Severson,D., Brown,S. and Knudson,D.
End sequencing of Aedes aegypti BACs
Unpublished (2003)
Other_GSSes: CSU-K34.113M15.T7
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: entae@tigr.org
Library was provided by Susan Brown and Dennis Knudson at Colorado
State University.
Seq primer: SP6
Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..773
                        /organism="Aedes aegypti"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:7159"
                        /clone="CSU-K34.113M15"
                        /clone_lib="CSU-K34"
                        /note="vector: pBACs3.6; Site_1: EcoRI; Source DNA: Aedes
                        aegypti; strain unknown [derived from freshly hatched
                        larvae at the Virus Research Centre, Poona, India.
                        Reference: SINGH, K. R. P., 1967 Cell cultures derived
                        from larvae of Aedes albopictus (Skuse) and Aedes aegypti
                        (L.). Current Science 36: 506-508]; ATC-10 cell line ATCC
                        CCL-125"

ORIGIN
Query Match               5.9%; Score 205; DB 8; Length 773;
Best Local Similarity    58.0%; Pred. No. 9.2e-31;
Matches 381; Conservative 0; Mismatches 275; Indels 1; Gaps 1;

Qy      2266 GACTCAGATTCAGATAGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACAGTGATCC 2325
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      767 GATTCAGATTAGATTCAGATTAGATTCAGATTAGATTCAGATTCAATTCAGATTTA 708
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      2326 GACTCAGACAGCGATTCCTGACTCCGACAGTGATTCGCGACTCAGACGCCGATTCGATTC 2385
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      707 GATTCAGATTGGGTCCGGATTAGATTCAGATTTAGATTCAGATTTAGATTCAGATTTG 648
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      2386 GACAGTGATTCGCGACTCAGTAGCGATTTCCGACTTCAGATCGCGACTCAGATTCAGACAGC 2445
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      647 GATTCAGATTAGATTCAGATTTAGATTCATATTTAGATTCAGATTTAGATTTCAATTTA 589
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      2446 GATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGATAGCGATTCAGATTCGCGACGACTCA 2505
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


LOCUS BX729701 699 bp mRNA linear EST 18-NOV-2003
DEFINITION BX729701 XGC-tadpole Xenopus tropicalis cDNA clone TtpA075e10 5',
RNA sequence.
ACCESSION BX729701
VERSION BX729701.1 GI:38402442
KEYWORDS EST.
SOURCE xenopus tropicalis (western clawed frog)
ORGANISM xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 699)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
JOURNAL Contact: Croning MDR
COMMENT Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TtpA075e10.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
cDNA was oligo dt primed from 5ug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..699
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TtpA075e10"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end"
ORIGIN
Query Match 4.9%; Score 171.8; DB 5; Length 699;
Best Local Similarity 53.3%; Pred. No. 4.5e-24;
Matches 362; Conservative 0; Mismatches 317; Indels 0; Gaps 0;
QY 1976 AGGATTCAGATTCCTGACCGAGTTTCAGATTCGGCAGCGATTCTTAATTCAGATAGCGGTT 2035
DB 20 AGGAATCCATAAAGTCACATGATGCTATTGGGGTCTCTGATCTGGTACTGGTACTGGTA 79
QY 2036 CAGATTCGGGTAGTGATTCATACATGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATT 2095
DB 80 CTGATACATGATACATGATCTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGT 139
QY 2096 CAGCAAGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGATTCCGACTCAGCGA 2155
DB 140 CTGGTTCTGGTACTGGTACTGATCTGGTACTGGTACTGGTACTGGTACTGGTACTGGTA 199
QY 2156 GCGATTCCGATCTGACAAATGACTCGGATTTCAGATTCAGATTCGATTCGACTCAGACAGTGACT 2215
DB 200 CTGATACATGGTCTCGGTACTGGTACTGATCTGGTACTGGTACTGGTACTGGTACTGGTA 259
QY 2216 CAGATTCGACAGTCACTCAGATTCAGATTCAGATTCGATTCGACTCAGACAGTCACTCAGATT 2275
DB 260 CTGGTACTGATCTGGTCTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTA 319
QY 2276 CAGATTCGATTCAGATTCAGATTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGACA 2335
DB 320 CTGGTACTGATCTGGTCTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTA 379
QY 2336 GCGATTTCGATTCGACAGTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2395

DB 380 CTGGTACTGATCTGGTACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 439
QY 2396 CCGACTCAGATAGCGATTCCGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATT 2455
DB 440 CTGATACATGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGG 499
QY 2456 CAGACAGCGATTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2515
DB 500 CTGGTACTGATCTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGATA 559
QY 2516 GTGACTCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATT 2575
DB 560 CTGATACATGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGATA 619
QY 2576 CAGACTCAGACAGTCACTCGGATTCAGCGAGTGAATTCGGATTCAGATTCAGATTCAGATTCAGATT 2635
DB 620 CTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGATA 679
QY 2636 CCGACAGTCACTCGGATTC 2654
DB 680 CTGGTACTGGTACTTATAC 698
RESULT 9
LOCUS CL129761/c 928 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-97D4 T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-97D4,
genomic survey sequence.
ACCESSION CL129761
VERSION CL129761.1 GI:40623396
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 928)
Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
AUTHORS A physical map of the xenopus tropicalis genome
TITLE Unpublished (2003)
JOURNAL Contact: Richard K Wilson
COMMENT Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 11
High quality sequence stop: 723.
FEATURES
Location/Qualifiers
1..928
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-97D4"
/clone_lib="ISB1"
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"
ORIGIN
Query Match 4.8%; Score 167.6; DB 9; Length 928;
Best Local Similarity 52.9%; Pred. No. 3.2e-23;
Matches 387; Conservative 0; Mismatches 339; Indels 6; Gaps 1;
QY 2135 CAACGGATTCGATTCAGCGAGCGATTCCGACTCAGACATGACTCGGATTCAGATTCAGATTCAG 2194
DB 750 CAATCTACCACTCGATACCTACTGAGTACTGAGTACTGAGTACTGAGTACTGAGTACTGAGTACT 691
QY 2195 ATTCTGACTCAGACAGTGAATTCAGATTCGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCTG 2254
DB 690 ACTGAGTACTACTGAGATTAATCTGAGATTAATCTGAGATTAATCTGAGTACTGAGTACTGAGTACT 631

2255 ACTCAGACAGTACTCAGATTTCAGATAGCAGATTTCAGATTCAGATTCAGATTCGCG 2314
Db |||||
630 ACTGAGTACCTACTCAGCTACCTACTGAGCTACCTACTGAGTACCTACTGAGTACCT 571
QY |||||
2315 ACAGTGATTCGCGACTCAGACAGCGATTCTGACTCCGACAGTGTATTCGCGACTCAGACAGCG 2374
Db |||||
570 ACTGAGTACCTACTCAGCTACCTACTGAGCTACCTACTGAGTACCTACTGAGTACCT 511
QY |||||
2375 ATTTCAGATTCGCGAGTGTATTCGCGACTCAGATAGCGATTTCGCGACTCAGATAGCGACTCAG 2434
Db |||||
510 ACTGAGTACCTACTCAGTGTATTCGCGACTCAGATAGCGATTTCGCGACTCAGATAGCGACTCAG 457
QY |||||
2435 ATTTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGATTCAGATTCGCG 2494
Db |||||
456 CTACCTACTGAGATTCCTACTGAGTACCTACTGAGTACCTACTGAGTACCTACTGAG 397
QY |||||
2495 ACAGTGACTCAGATTCGCGAGTGTATTCGCGACTCAGATTCAGATTCAGATTCGCGAGTGT 2554
Db |||||
396 ATACCTACTGAGAACTTACGGAGATAAATCTACTGAGTACCTACTGAGTACCTACTGAG 337
QY |||||
2555 ACTCAGATTCGCGAGTGTATTCGCGACTCAGATTCAGATTCAGATTCAGATTCGCG 2614
Db |||||
336 TTACCTACTGATTCCTACTGAGTACCTACTGAGTACCTACTGAGTACCTACTGAGTACCT 277
QY |||||
2615 ATTTCAGATTCGCGAGTGTATTCGCGACTCAGATTCAGATTCAGATTCAGATTCGCG 2674
Db |||||
276 ATACCTACTGAGATTCCTACTGAGTACCTACTGAGTACCTACTGAGTACCTACTGAG 217
QY |||||
2675 ATACCGACTCAGATTCAGATTCAGATTCGCGACTCAGATTCAGATTCAGATTCAGATTCGCG 2734
Db |||||
216 ATACCTACTGAGTACCTACTGAGTACCTACTGAGTACCTACTGAGTACCTACTGAGTACCT 157
QY |||||
2735 ATTTCAGATTCGCGAGTGTATTCGCGACTCAGATTCAGATTCAGATTCAGATTCGCG 2794
Db |||||
156 ATAACTACTGAGATTCCTACTGAGTACCTACTGAGTACCTACTGAGTACCTACTGAGTACCT 97
QY |||||
2795 ATAGTACTCAGATTCAGATTCAGATTCGCGACTCAGATTCAGATTCAGATTCAGATTCGCG 2854
Db |||||
96 ATACCTACTGAGATTCCTACTGAGTACCTACTGAGTACCTACTGAGTACCTACTGAGTACCT 37
QY |||||
2855 ATTTCGAGTACG 2866
Db |||||
36 ATACCTACTGAG 25

CN689123 563 bp mRNA linear EST 17-MAY-2004
E0270H03-5 NIA Mouse Embryonic Stem (ES) cell (Lif-, 48 h, high density) cDNA library (Long) Mus musculus cDNA clone NIA:E0270H03 IMAGE:30856118 5', mRNA sequence.

CN689123
CN689123.1 GI:47455569
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 563)
Sharov, A.A., Piao, Y., Matoba, R., Dudekula, D.B., Qian, Y., VanBuren, V., Falco, G., Martin, P.R., Stagg, C.A., Bassey, U.C., Wang, Y., Carter, M.G., Hamatani, T., Aiba, K., Akutsu, H., Sharova, L., Tanaka, T.S., Kimber, W.L., Yoshikawa, T., Jaradat, S.A., Pantano, S., Nagaraja, R., Boheler, K.R., Taub, D., Hodes, R.J., Longo, D.L., Schlagesinger, D., Keller, J., Klotz, E., Kelsae, G., Umezawa, A., Vescovi, A.L., Rossant, J., Kunath, T., Hogan, B.L., Curci, A., D'Urso, M., Kelsae, J., Hide, W. and Ko, M.S.
Transcriptome analysis of mouse stem cells and early embryos
PLoS Biol. 1 (3), 410-419 (2003)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health

FEATURES
Source

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: E0270 row: H column: 03
Seq primer: M13 Reverse
High quality sequence stop: 563
POLYA=No.

Location/Qualifiers
1. 563
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129Sv/EvTac"
/db_xref="niaEST:E0270H03-5"
/db_xref="taxon:10090"
/clone="NIA:E0270H03 IMAGE:30856118"
/sex="Male"
/tissue_type="Embryonic Stem Cell"
/cell_line="129.3 ES cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Embryonic Stem (ES) cell (Lif-, 48 h, high density) cDNA library (Long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library [Ref. Genome Res. 11: 1553-1558 (2001). PMID: 11544199]. ES cells were plated at density 3x104/cm2 on gelatin-coated plates and cultured for 48 hrs at 37 OC, 5% CO2. Culture medium: DMEM supplemented with 15% FBS, 2 mM L-glutamine, 0.1 mM NEAA, 1mM Sodium pyruvate, 0.1 mM beta-mercaptoethanol, 100 U/ml penicillin, and 100 ug/ml streptomycin. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]. 5'-pGACTAGTCTCAGATTCGAGCGCGCCCTTTTTCCTTTT-3' from 2ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.4 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 4.8%; Score 167.4; DB 7; Length 563;
Best Local Similarity 56.1%; Pred. No. 3.4e-23;
Matches 315; Conservative 0; Mismatches 246; Indels 0; Gaps 0;

QY 2131 TCAGCAAGCGATTCCGAGTCTCAGCGAGCGATTCCGAGTCTCAGCAATGACTCGGATTCAGAT 2190
Db |||||
563 TCACCTGGGACTTGACCTCAGCTGGAGATTGGCCCTCAGCTGGTGTGACTTGACCTCAGCT 504
QY 2191 AGCGATTCTGACTCAGACAGTGTACTCAGATTCCGACAGTGTACTCAGATTCAGATTCAGAT 2250
Db |||||
503 GGAGATTGGCCCTCAGCTGGTGTGACTTGCCCTCAGCTGGAGACTTGGCCTCACCTGGTGAC 444
QY 2251 TCTGACTCAGACAGTGTACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2310
Db |||||
443 TTGACCTCAGCTGGAGATTGGCCCTCAGCTGGAGATTGGCCCTCAGCTGGAGATTGGCT 384
QY 2311 TCCGACAGTGTATTCGAGTCTCAGACAGCGATTCTGACTCCGACAGTGTATTCGAGTCCGACTCAGAC 2370
Db |||||
383 TCAGCTGGAGATTGGCCCTCAGATGTGACTTGGCCCTCAGCTGGAGACTTCACTGTAGCT 324
QY 2371 AGCGATTCTGATTCCGACAGTGTATTCGAGTCTCAGATTCAGATTCAGATTCAGATTCAGAT 2430
Db |||||
323 GGAGATTGGCTTCAGCTGTGACTTGGCTTCAGCTGGTGTGACTTGGCCTCAGCTGGAGAT 264

LOCUS AL162088 1526 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp761l11023_r1.761 (synonym: hamy2) Homo sapiens cDNA clone
DKFZp761l11023_5', mRNA sequence.
ACCESSION AL162088
VERSION AL162088.1 GI:7330401
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1526)
AUTHORS Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GfR (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
SI sequence also available.
This clone (DKFZp761l1023) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..1526
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp761l1023"
/tissue_type="amygdala"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="761 (synonym: hamy2)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 4.6%; Score 160.2; DB 1; Length 1526;
Best Local Similarity 62.4%; Pred. No. 1e-21;
Matches 285; Conservative 0; Mismatches 168; Indels 4; Gaps 2;
QY 2354 GTGATTCGGACTCAGACGCGATTCCGACAGTGATTCGGACTCAGATTCGCGATT 2413
DB 1526 GAGATAGACACACAGATAGAGATAGATAGATAGATAGAT--AGATAGAGATAGAGATA 1469
QY 2414 CCGACTCAGATAGCGACTCAGATTCCAGACGCGATTCCAGATTCAGACAGCGATTCCAGATT 2473
DB 1468 GAGACACAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGACACA 1409
QY 2474 CAGATACGCGATTCCAGATTCGG--ACAGTGACTCAGATTCGGACAGTGACTCGGATTCAGA 2531
DB 1408 CAGATAGATAGATAGATAGATATAGATAGATAGACACACAGATAGATAGATAGACAGA 1349
QY 2532 TAGCGATTCCAGATTCGACAGTGACTCAGATTCCGACAGTGACTCAGACTCAGACAGTGA 2591
DB 1348 TAGAGATAGACACACAGATAGATAGATAGATAGATAGATAGATAGATAGATAGACACAGA 1289
QY 2592 TTCGATTCCAGCGATTCGGATTCCGATTCCGATTCCGATTCCGACTCCGACAGTGACTCGGA 2651
DB 1288 TAGAGATAGATAGATAGATAGACACAGATAGATAGATAGATAGATAGATAGACACAGA 1229
QY 2652 TTCAGATAGCGACTCCAGACTCCGATTCGCGATTCCAGATAGCGACTCCGACTCAGA 2711
DB 1228 TAGAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGACACACAGA 1169
QY 2712 TAGCGATTCCAGATTCAGACAGCGATTCCAGATTCAGACAGCGATTCCAGATTCAGACAGCGA 2771
DB 1168 TAGAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAGA 1109

QY 2772 CTCAGACAGTGACTCAGATTCCAGATTCAGATAGTAGTACTCGGAT 2808
DB 1108 TAGAGATAGACACACAGATAGATAGATAGATAGATAGAT 1072

Search completed: November 5, 2004, 12:54:47
Job time : 10227 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:48:10 ; Search time 176.713 Seconds
(without alignments)
1894.005 Million cell updates/sec

Title: US-09-679-643-2

Perfect score: 4725

Sequence: 1 MNMKKEKHAIRKKSIGVAS.....LASIGSLLFRKKENKDKK 933

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_23Sep04.*
1: geneseq1980s.*
2: geneseq1990s.*
3: geneseq2000s.*
4: geneseq2001s.*
5: geneseq2002s.*
6: geneseq2003as.*
7: geneseq2003bs.*
8: geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4725	100.0	933	3 AAY58435	Aay58435 Staphyloc
2	4725	100.0	933	4 AAB69508	Aab69508 Staphyloc
3	4717	99.8	933	6 ABJ18947	Abj18947 Pathogen
4	4662	98.7	927	6 ABM72221	Abm72221 Staphyloc
5	4662	98.7	936	2 AAW89801	Aaw89801 Staphyloc
6	4359	92.3	935	6 ABU16402	Abu16402 Protein e
7	4357	92.2	1021	4 AAU33975	Aau33975 Staphyloc
8	4357	92.2	1021	4 AAU36951	Aau36951 Staphyloc
9	2671	56.5	520	5 AAE29262	Aae29262 Staphyloc
10	2514	53.2	496	5 AAU75490	Aau75490 S. aureus
11	2514	53.2	496	6 ADA89664	Ada89664 Staphyloc
12	2514	53.2	496	6 ADA89663	Ada89663 Staphyloc
13	1813.5	38.4	913	6 ABJ18917	Abj18917 Pathogen
14	1808.5	38.3	918	2 AAY08640	Aay08640 S. aureus
15	1727	36.6	331	5 AAE29263	Aae29263 Staphyloc
16	1727	36.6	345	2 AAW31555	Aaw31555 Fibronect
17	1654.5	35.0	877	6 ADA89539	Ada89539 Staphyloc
18	1654.5	35.0	877	6 ABM72702	Abm72702 Staphyloc
19	1638.5	34.7	877	6 ABU42504	Abu42504 Protein e
20	1628	34.5	1092	2 AAW41602	Aaw41602 Staphyloc
21	1628	34.5	1092	7 ABM79019	Abm79019 Staphyloc
22	1607.5	34.0	1633	3 ABU42513	Abu42513 Protein e
23	1584.5	33.5	1802	3 AAY83170	Aay83170 Cell wall
24	1584.5	33.5	1802	3 AAY70119	Aay70119 Staph. ep
25	1552	32.8	1920	6 ABU43489	Abu43489 Protein e

ALIGNMENTS

RESULT 1

AAV58435

ID AAY58435 standard; protein; 933 AA.

XX AC

XX AAY58435;

XX AC

XX 27-MAR-2000 (first entry)

XX AC

XX DE Staphylococcus aureus fibrinogen binding ClfA protein.

XX ClfA; fibrinogen binding protein; bacterial colonisation;

XX KW indwelling medical device; staphylococcal infection.

XX OS Staphylococcus aureus.

XX XX

XX Key Location/Qualifiers

XX Peptide 1..39

XX /note= "Signal peptide"

XX Region 40..559

XX /note= "Region A"

XX Region 332..550

XX /note= "Fibrinogen-binding region"

XX Region 560..867

XX /note= "Region R"

XX Region 896..900

XX /note= "Gram positive wall-associated consensus motif"

XX US6008341-A.

XX 28-DEC-1999.

XX 22-AUG-1994; 94US-00293728.

XX 22-AUG-1994; 94US-00293728.

XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

XX Foster TJ, Mcdevitt DL;

XX WPI; 2000-096389/08.

XX N-PSDB; AAZ55832.

XX Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus

XX aureus, useful for treatment or prevention of infections.

XX Disclosure; Fig 2A-1-4; 35pp; English.

XX This sequence represents the Staphylococcus aureus fibrinogen-binding

XX CC

app. cat

protein, ClfA. ClfA is an important receptor involved in *S. aureus* colonisation of indwelling medical devices (e.g., catheters, artificial heart valves). Shortly after implantation, the surfaces of medical devices become coated with host plasma and matrix proteins such as fibrinogen and fibronectin, and there is considerable evidence to suggest that bacterial adherence to fibrinogen/fibrin is important in the initiation of device-related infection. The fibrinogen-binding region of ClfA is thought to reside between residues 332 and 550 in a region designated A. The protein also contains a repeated region (region R) comprising 154 repeats of the dipeptide Ser-Asp, and the C-terminus contains features present in surface proteins of other Gram positive bacteria that are responsible for anchoring the protein to the cell wall and cell membrane. ClfA, or its fragments, may be used to block *S. aureus* colonisation of wounds, to prevent adherence of *S. aureus* to indwelling medical devices, as vaccines to protect against *S. aureus* infection (e.g., mastitis in ruminants), to raise specific antibodies, and for diagnosis (by agglutination or immunoassay). The specific antibodies are used for passive immunisation, to block infection of wounds or adhesion of *S. aureus* and for diagnosis. Nucleotides encoding ClfA and its fragments may be used as diagnostic probes.

Sequence 933 AA;

Query Match	100.0%	Score 4725;	DB 3;	Length 933;
Best Local Similarity	100.0%;	Pred. No. 5.8e-233;		
Matches 933;	Conservative 0;	Mismatches 0;	Indels 0	

[illegible][illegible]

RESULT 2

RESOLUT 2
AAB69508
ID AAB69508 standard; protein: 933 AA.

XX AAB69508;

XX DT 23-APR-2001 (first entry)

Staphylococcus aureus Clfa

Staphylococcus aureus: clfa: antibi

OS Staphylococcus aureus
XX
XX
KW Fibrinogen binding protein; bacterial infection; mastitis
XX

XX
DNI IIS6177084-B1

XX 02 53M 0001

XX
10 OCT 1980
DE

XX
DE 22 MAY 1964
0410 0000700

XX

XX

DR N-PSDB; AAF58593.

PT Novel Staphylococ

PT indwelling medical devices and for diagnosing staphylococcus aureus PT infection.

PS Claim 5; Fig 2; 30pp; English.

CC The present sequence is a novel *Staphylococcus aureus* fibrinogen binding
CC protein. It is useful as a vaccine to protect against human and animal
CC infections caused by *S. aureus*, such as against mastitis, to block
CC *S. aureus* from colonising and infecting a wound, to block adherence of
CC *S. aureus* to indwelling medical devices such as catheters, replacement
CC heart valves and cardiac assist devices. The protein can be used to
CC diagnose bacterial infections. Polyclonal and monoclonal antibodies
CC raised against the fibrinogen binding protein are useful for passively
CC immunising against infections caused by *S. aureus*, to prevent infection of
CC a wound and to diagnose bacterial infections

Sequence 933 AA;

Query Match 100.0%; Score 4725; DB 4; Length 933;

Best Local Similarity
100.0%; Pred. No. 5.8e-233;

Matches	933	Conservative	0	Mismatches	0	Indels	0	Gaps	0
---------	-----	--------------	---	------------	---	--------	---	------	---

oppl

QY 1 MNMKEKHAIRKKSIGVASVLVGTLLIGFGLLSKEADASENSVTQSDASNEKSDSS 60
DB 1 MNMKEKHAIRKKSIGVASVLVGTLLIGFGLLSKEADASENSVTQSDASNEKSDSS 60
QY 61 SVSAAPEKTDVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEPTVTEATT 120
DB 61 SVSAAPEKTDVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEPTVTEATT 120
QY 121 TTQANTPATQSSNTNAEELVNOTSNFTFNDTNTVSSVNSPQNSNAENVSTQDST 180
DB 121 TTQANTPATQSSNTNAEELVNOTSNFTFNDTNTVSSVNSPQNSNAENVSTQDST 180
QY 181 EATPSNNEAPQSTDSAKDQVNVQAVNTSAPRMAFSLAAVAADAPAGTDITNQLNVT 240
DB 181 EATPSNNEAPQSTDSAKDQVNVQAVNTSAPRMAFSLAAVAADAPAGTDITNQLNVT 240
QY 241 VGDSDGTTVPHQAGYVKNLNGFVSPNSAVKGDTEKITVPEKELNNGVTSKAVPPIMAG 300
DB 241 VGDSDGTTVPHQAGYVKNLNGFVSPNSAVKGDTEKITVPEKELNNGVTSKAVPPIMAG 300
QY 301 DQVLAVGVIDSGNVITFTDYVNTKDDVKAATLMPAYIDPENVKKTGNVTLATIGSTT 360
DB 301 DQVLAVGVIDSGNVITFTDYVNTKDDVKAATLMPAYIDPENVKKTGNVTLATIGSTT 360
QY 361 ANKTVLVDYKGYKFNLSIKGTIDQDKTNTYRQTIYVNPNGDNVIAPLVGNLKPNT 420
DB 361 ANKTVLVDYKGYKFNLSIKGTIDQDKTNTYRQTIYVNPNGDNVIAPLVGNLKPNT 420
QY 421 DSNALIDQNTSIKVKVDNAADLSESYFVNPENFEDVTSVNIITFPNPQYKVEFNTPD 480
DB 421 DSNALIDQNTSIKVKVDNAADLSESYFVNPENFEDVTSVNIITFPNPQYKVEFNTPD 480
QY 481 DQITTPYVNVGHIDPNSGDLALRSLTYGNSNIIWRMSWDNEVAFNNGSGSGDID 540
DB 481 DQITTPYVNVGHIDPNSGDLALRSLTYGNSNIIWRMSWDNEVAFNNGSGSGDID 540
QY 541 KPVPPEQDPGELEPIPEPDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSD 600
DB 541 KPVPPEQDPGELEPIPEPDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSD 600
QY 601 SDSDSASDSDSASD 660
DB 601 SDSDSASDSDSASD 660
QY 661 SD 720
DB 661 SD 720
QY 721 SD 780
DB 721 SD 780
QY 781 SD 840
DB 781 SD 840
QY 841 SD 900
DB 841 SD 900
QY 901 SEDEANTSLIWGLLASIGSLLLFRKKENKDKK 933
DB 901 SEDEANTSLIWGLLASIGSLLLFRKKENKDKK 933

RESULT 3

ABJ18947

ID ABJ18947 standard; protein; 933 AA.

XX ABJ18947;

AC ABJ18947;

DT 06-MAR-2003 (first entry)

XX

DE Pathogen specific antigen related staphylococcal protein SEQ ID No 93.
XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW autoimmune disease; HIV; hepatitis.
XX Staphylococcus sp.
XX WO200259148-A2.
XX 01-AUG-2002.
XX 21-JAN-2002; 2002WO-EP000546.
XX 26-JAN-2001; 2001AT-00000130.
XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX Meinke A, Nagy E, Von Ahsen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vytvytska O, Btz H, Dryla A, Weichhart T, Hafner M;
PI Tempelmaier B;
XX WPI; 2003-075410/07.
XX Identifying, isolating and producing hyperimmune serum-reactive antigens
XX from a pathogen, for preparing vaccine or medicament for treating or
XX preventing e.g. staphylococcal infections, comprises providing antibody
XX preparation.

Example 7; Page 160; 252pp; English.

XX The invention relates to a novel method for identifying, isolating and
XX producing hyperimmune serum-reactive antigens from a pathogen, tumour,
XX allergen, a tissue or host prone to auto-immunity, where the antigens are
XX used in a vaccine, comprises providing antibody preparation from a plasma
XX pool of a type of animal, or individual sera with antibodies against the
XX specific pathogen, tumour, allergen, tissue or host prone to auto-
XX immunity. The hyperimmune serum-reactive antigens comprising any of the
XX 62 sequences of 53-2261 amino acids fully defined in the specification,
XX or their hyperimmune fragments are useful for the manufacture of a
XX pharmaceutical preparation, particularly a vaccine against staphylococcal
XX infections or colonisation against S. aureus or S. epidermidis. The
XX preparation of antibodies is useful for the manufacture of a medicament
XX for treating or preventing staphylococcal infections or colonisation
XX against S. aureus or S. epidermidis. The antibody preparations may also
XX be used for diagnostic and imaging purposes. Other conditions that can be
XX treated include cancer, autoimmune diseases or infections caused by viral
XX (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
XX sequence represents a staphylococcal protein relating to the method for
XX identifying and producing pathogen specific antigens of the invention

SQ Sequence 933 AA;

Query Match

Best Local Similarity 99.8%; Score 4717; DB 6; Length 933;

Matches 932; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNMKEKHAIRKKSIGVASVLVGTLLIGFGLLSKEADASENSVTQSDASNEKSDSS 60
DB 1 MNMKEKHAIRKKSIGVASVLVGTLLIGFGLLSKEADASENSVTQSDASNEKSDSS 60
QY 61 SVSAAPEKTDVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEPTVTEATT 120
DB 61 SVSAAPEKTDVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEPTVTEATT 120
QY 121 TTQANTPATQSSNTNAEELVNOTSNFTFNDTNTVSSVNSPQNSNAENVSTQDST 180
DB 121 TTQANTPATQSSNTNAEELVNOTSNFTFNDTNTVSSVNSPQNSNAENVSTQDST 180
QY 181 EATPSNNEAPQSTDSAKDQVNVQAVNTSAPRMAFSLAAVAADAPAGTDITNQLNVT 240
DB 181 EATPSNNEAPQSTDSAKDQVNVQAVNTSAPRMAFSLAAVAADAPAGTDITNQLNVT 240

QY 241 VGDSTTVYPHQAGYVVKLYGFSVPNSAVKGDTEKITVPKELNLNGVTSTAKVPPIMAG 300
DB 241 VGDSTTVYPHQAGYVVKLYGFSVPNSAVKGDTEKITVPKELNLNGVTSTAKVPPIMAG 300
QY 301 DQVLAVGVDSDGNVYITFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360
DB 301 DQVLAVGVDSDGNVYITFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360
QY 361 ANKTVLVDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYVNPFGDNVIAPIVLTGNLKPT 420
DB 361 ANKTVLVDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYVNPFGDNVIAPIVLTGNLKPT 420
QY 421 DSNALIDQNTSIVKVKVDNAADLSSEYFVNPENFEDVNTSNVNIFFPNPQYKVFENTPD 480
DB 421 DSNALIDQNTSIVKVKVDNAADLSSEYFVNPENFEDVNTSNVNIFFPNPQYKVFENTPD 480
QY 481 DQITTPYIVVNGHIDPNKGDALRSTLYGYNSNIWRSMSWDNEVAFNNGSGGID 540
DB 481 DQITTPYIVVNGHIDPNKGDALRSTLYGYNSNIWRSMSWDNEVAFNNGSGGID 540
QY 541 KPVVPEQDPGEIEPIPEDSDPGSDGSDNSDGSSTSDSGSDSASDSASA 600
DB 541 KPVVPEQDPGEIEPIPEDSDPGSDGSDNSDGSSTSDSGSDSASDSASA 600
QY 601 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 660
DB 601 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 660
QY 661 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 720
DB 661 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 720
QY 721 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 780
DB 721 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 780
QY 781 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 840
DB 781 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 840
QY 841 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 900
DB 841 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 900
QY 901 SEDEANTSLIWGLLASIGSLLLFRKKENKOKK 933
DB 901 SEDEANTSLIWGLLASIGSLLLFRKKENKOKK 933

RESULT 4
ID ABM72221
XX ABM72221 standard; protein; 927 AA.
AC ABM72221;
XX 20-NOV-2003 (first entry)
XX Staphylococcus aureus protein #1461.
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX enzymatic assay; antibiotic target.
XX Staphylococcus aureus.
XX WO200204868-A2.
XX 28-NOV-2002.
XX 27-MAR-2002; 2002WO-IB002637.
XX 27-MAR-2001; 2001GB-00007661.
XX

PA (CHIR-) CHIRON SPA.
XX Masignani V, Mora M, Scarselli M;
XX WPI; 2003-120786/11.
DR N-PSDB; ACF73781.
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX Claim 1; SEQ ID NO 2922; 49pp; English.
XX The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus proteins of the invention
XX Sequence 927 AA;
Query Match 98.7%; Score 4662; DB 6; Length 927;
Best Local Similarity 99.0%; Pred. No. 9.4e-230;
Matches 924; Conservative 0; Mismatches 3; Indels 6; Gaps 1;
QY 1 MMKKKEKHAIRKKSIGVASVLVGLIGLLSSKEADASENSVTQSDASNESKSDSS 60
DB 1 MMKKKEKHAIRKKSIGVASVLVGLIGLLSSKEADASENSVTQSDASNESKSDSS 60
QY 61 SVSAAKPTDDTNVSDTKTSNTNGETSVAQNPACQETTSSTNATTEETPTVTEATT 120
DB 61 SVSAAKPTDDTNVSDTKTSNTNGETSVAQNPACQETTSSTNATTEETPTVTEATT 120
QY 121 TTQANTPATTQSSNTNAEELVNQTSNETTNDNTVSVSPQNSTNAENVSTTQDTST 180
DB 121 TTQANTPATTQSSNTNAEELVNQTSNETTNDNTVSVSPQNSTNAENVSTTQDTST 180
QY 181 EATPSNESAPQSITDASNDKVVQAVNTSAPRMRAFSLAAVAADAPAGTITNLTNT 240
DB 181 EATPSNESAPQSITDASNDKVVQAVNTSAPRMRAFSLAAVAADAPAGTITNLTNT 240
QY 241 VGIDSGTTVYPHQAGYVVKLYGFSVPNSAVKGDTEKITVPKELNLNGVTSTAKVPPIMAG 300
DB 241 VGIDSGTTVYPHQAGYVVKLYGFSVPNSAVKGDTEKITVPKELNLNGVTSTAKVPPIMAG 300
QY 301 DOVLAVGVDSDGNVYITFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360
DB 301 DOVLAVGVDSDGNVYITFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360
QY 361 ANKTVLVDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYVNPFGDNVIAPIVLTGNLKPT 420
DB 361 ANKTVLVDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYVNPFGDNVIAPIVLTGNLKPT 420
QY 421 DSNALIDQNTSIVKVKVDNAADLSSEYFVNPENFEDVNTSNVNIFFPNPQYKVFENTPD 480
DB 421 DSNALIDQNTSIVKVKVDNAADLSSEYFVNPENFEDVNTSNVNIFFPNPQYKVFENTPD 480
QY 481 DQITTPYIVVNGHIDPNKGDALRSTLYGYNSNIWRSMSWDNEVAFNNGSGGID 540
DB 481 DQITTPYIVVNGHIDPNKGDALRSTLYGYNSNIWRSMSWDNEVAFNNGSGGID 540
QY 541 KPVVPEQDPGEIEPIPEDSDPGSDGSDNSDGSSTSDSGSDSASDSASA 600
DB 541 KPVVPEQDPGEIEPIPEDSDPGSDGSDNSDGSSTSDSGSDSASDSASA 600
QY 601 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 660
DB 601 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 660

Dd 904 SEDEANTSLIWGLLASIGSLLFRKKNKDKK 936

RESULT 6

ABU16402
ID ABU16402 standard; protein; 935 AA.

XX AC ABU16402;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #1929.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Staphylococcus aureus.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 23-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KJ, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA20272.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 44326; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 935 AA;

Query Match

Best Local Similarity 92.3%; Score 4359; DB 6; Length 935;

Matches 865; Conservative 28; Mismatches 40; Indels 2; Gaps 1;

QY 1 MMKKKEKHAIRKKSIGVASVLVGLTIGFGLSSKEADASENSVTOSDSASNEKNDSS 60

Db 1 MMKKKEKHAIRKKSIGVASVLVGLTIGFGLSSKEADASENSVTOSDSASNEKNDSS 60

QY 61 SVSAAFKTDDTNVSDTKTSSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGATT 120

Db 61 SVSAAFKTDDTNVSDTKTSSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGATT 120

QY 121 TTQANTPATTOSSNTNABELVQNSNETTNDNTVSSVNSPQNSNAENVSTTQDTST 180

Db 121 TTQANTPATTOSSNTNABELVQNSNETTNDNTVSSVNSPQNSNAENVSTTQDTST 180

QY 181 EATPSNNEAPOSTDASNKDVVQAVNTSAPSWRAFSLAAVAADAPAACTDITNOLTNVT 240

Db 181 EATPSNNEAPONTDASNKDVVQAVNTPRNPAPSLAAVAADAPAACTDITNOLTNVT 240

QY 241 VGIDSGTTVYPHQAGYVKLYNGFSVPNSAVKGDTFKITVPKELNLNGVTSTAKVPPIMAG 300

Db 241 VTIDSGTTVYPHQAGYVKLYNGFSVPNSAVKGDTFKITVPKELNLNGVTSTAKVPPIMAG 300

QY 301 DOVLANGVIDSGNVIIYFTDYVNTKDDVKATLTPAYIDPENVKKTGNVTIATGIGSTT 360

Db 301 DOVLANGVIDSGNVIIYFTDYVNDKENVTANITPAYIDPENVTIATGIGSTT 360

QY 361 ANKTVLVDEYKQKFNLSIKGTIQQDKNTTYQTIVNPSGDNVAPLVTGNLKPN 420

Db 361 ASKTVLVDEYKQGFHNLSIKGTIQQDKNTTYQTIVNPSGDNVAPLVTGNLKPN 420

QY 421 DSNALIDQNTSIIKVKVDNAADLSSEYFVNPFEDVTNSVNIITFPNPNQYKVFNTPD 480

Db 421 KSNALIDAKNTDIKVYVDNANDLSSEYVNPSPDEFVINOVRISFPNPNQYKVFNTPD 480

QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGNSNIIWRSMWNEVAFNNGSGSGDGD 540

Db 481 DOITTPYIVVNGHIDPASTGDLALRSTFYGDSNFIWRSMWNEVAFNNGSGSGDGD 540

QY 541 KPVPPEQDEPGEIPIPEDSDPGSGSDNSDSDSDSDSDSDSDSDSDSDSDSDSDSA 600

Db 541 KPVPPEQDEPGEIPIPEDSDPGSGSDNSDSDSDSDSDSDSDSDSDSDSDSDSDSA 600

QY 601 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 660

Db 601 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 660

QY 661 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 720

Db 661 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 720

QY 721 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 780

Db 721 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 780

QY 781 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 838

Db 781 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 840

QY 839 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 898

Db 841 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 900

QY 899 TGSDEANTSLIWGLLASIGSLLFRKKNKDKK 933

Db 901 TGSDEANTSLIWGLLASIGSLLFRKKNKDKK 935

DT 14-FEB-2002 (first entry)
 XX Staphylococcus aureus cellular proliferation protein #1121.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS54810.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Example 3; SEQ ID NO 12544; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1021 AA;

Query Match 92.2%; Score 4357; DB 4; Length 1021;
 Best Local Similarity 85.5%; Pred. No. 3.8e-214;
 Matches 873; Conservative 27; Mismatches 33; Indels 88; Gaps 3;

QY 1 MNMKKEKHARKKSIGVASVLVGLTIGFLLSKSEADASENSVTQSDASNESKSDSS 60
 DB 1 MNMKKEKHARKKSIGVASVLVGLTIGFLLSKSEADASENSVTQSDASNESKSDSS 60
 QY 61 SVSAAPKTDITNVSDTKTSNTNNGETSVAPNPAQOETTQSSSTNATTEPTVGEATT 120
 DB 61 SVNAPKTDITNVSDTKTSNTNNGETSVAPNPAQOETTQSSSTNATTEPTVGEATT 120
 QY 121 TTQANTPATQSSNTNNAELVNQTSNETTFDNTVSSVNSPQNSTNAENVSTQDTST 180
 DB 121 ATKQANTPATQSSNTNNAELVNQTSNETASNDINTVSSVNSPQNSTNAENVSTQDTST 180

QY 181 EATPSNNEAPQSTDAKNQVNVQAVNTSAPRMRAFSLAAVAADAPAGTDITNQLTNVT 240
 DB 181 EATPSNNEAPQSTDAKNQVNVQAVNTSAPRMRAFSLAAVAADAPAGTDITNQLTDVK 240
 QY 241 VGIDSGTTVYVPHQAGYVVKLYGFSVPNSAVKGDFTKITVPKELNGLNGVTSTAKVPPIMAG 300
 DB 241 VTIDSGTTVYVPHQAGYVVKLYGFSVPNSAVKGDFTKITVPKELNGLNGVTSTAKVPPIMVG 300
 QY 301 DOVLANGVIDSGNVITYTFTDYNTKDYKATILTMAYIDPENVKITGNVLTATGSGTT 360
 DB 301 DOVLANGVIDSGNVITYTFTDYNTKDYKATILTMAYIDPENVKITGNVLTATGSGTT 360
 QY 361 ANKTLVDYKCYKGFYNLSIKGTIDQIDKTNNTYROTIIYVNPSGDNVIAPLVTGNLKPN 420
 DB 361 ASKTLVDYKCYKGFYNLSIKGTIDQIDKTNNTYROTIIYVNPSGDNVIAPLVTGNLIPNT 420
 QY 421 DSNALIDQONTSIKVYKVDNAADLSSEYFVNPENFEDVTNSVNIITPPNQYKVEFTED 480
 DB 421 KGNALIDAKNTDIKVYRVNDANDLSSEYVYVNPEDFVTNQVRISFPNANQYKVEFTDD 480
 QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYNSNLIWESMWDNEVAFNNGSGSDGID 540
 DB 481 DOITTPYIVVNGHIDPASTGDLALRSTFYGYDSNFIWESMWDNEVAFNNGSGSDGID 540
 QY 541 KFWPEQPDPEGEIPIPEDSDSDPGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSG 600
 DB 541 KFWPEQPDPEGEIPIPEDSDSDPGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSG 600
 QY 601 SDDSASDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 622
 DB 601 SDDSASDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 660
 QY 623 SNDSD 682
 DB 661 SDDSASDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 720
 QY 683 SDDSASDSDSGSD 742
 DB 721 SDDSASDSDSGSD 780
 QY 743 SDDSASDSDSGSD 802
 DB 781 SDDSASDSDSGSD 840
 QY 803 SDDSASDSDSGSD 832
 DB 841 SDDSASDSDSGSD 900
 QY 833 --SDDSASDSDSGSDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 872
 DB 901 SGSDSDSGSDSDSGSDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 960
 QY 873 PNPSPKNGTASNKNEAKDSKEPLPDTGSEDEANTSLIWGLLASTIGSLLLFRKKEKDK 932
 DB 961 PNPSPKNGTASNKNEAKDSKEPLPDTGSEDEANTSLIWGLLASTIGSLLLFRKKEKDK 1020
 QY 933 K 933
 DB 1021 K 1021

RESULT 9
 AAE29262
 ID AAE29262 standard; protein; 520 AA.
 XX
 AC AAE29262;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Staphylococcus aureus Clf40 protein.
 XX
 KW Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
 immunological; staphylococcal infection; impetigo; pneumonia; furuncle;

KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.

XX Staphylococcus aureus.

OS

XX

XX

PN WO200272600-A2.

PD

PD 19-SEP-2002.

PF

PF 28-JAN-2002; 2002WO-US002296.

XX

XX 26-JAN-2001; 2001US-0264072P.

PR 12-MAR-2001; 2001US-0274611P.

PR 18-JUN-2001; 2001US-0298413P.

PR 30-JUL-2001; 2001US-0308116P.

XX

XX (INH1-) INHIBITEX INC.

XX

XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;

PI WPI; 2002-759834/82.

XX N-PSDB; AAD46861.

DR

DR

XX

XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for

PT treating or preventing Staphylococcus aureus infection e.g. wound

PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in

PT a human or animal.

XX

XX Claim 9; Page 67-69; 80pp; English.

PS

XX

XX The invention relates to monoclonal antibody which binds the clumping

CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA

CC monoclonal antibody is useful for treating or preventing S. aureus

CC infection in a human or animal, and for inhibiting the binding of

CC Staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment

CC S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3

CC protein is useful for inducing an immunological response in a human or

CC animal. These staphylococcal infections include wound infections, sepsis,

CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The

CC present sequence is Staphylococcus aureus Clf40 protein

XX

SQ Sequence 520 AA;

Query Match 56.5%; Score 2671; DB 5; Length 520;

Best Local Similarity 100.0%; Pred. No. 1.9e-128; Indels 0; Gaps 0;

Matches 520; Conservative 0; Mismatches 0;

40 SENSVTQSDSASNESKNDSSVSAAPKTDITNNVSDTKTSSNTNNGTSSVAQNPAQQT 99

1 SENSVTQSDSASNESKNDSSVSAAPKTDITNNVSDTKTSSNTNNGTSSVAQNPAQQT 60

100 QSSSTNATTEPTVGEATTTTNOANTPATQSSNTNNAELVNOTSNETTNDNTVSS 159

61 QSSSTNATTEPTVGEATTTTNOANTPATQSSNTNNAELVNOTSNETTNDNTVSS 120

160 VNSPQNSTNAENVSTQDTSSTATSNESAPOSQSDASKNVQAVNTSAPRMAFSLA 219

121 VNSPQNSTNAENVSTQDTSSTATSNESAPOSQSDASKNVQAVNTSAPRMAFSLA 180

220 AVAADAPAGATDTNLTNTVTVGIDSGTTVYVPHQAGYVKNLYGFSVPNSAVKGDTPKTI 279

181 AVAADAPAGATDTNLTNTVTVGIDSGTTVYVPHQAGYVKNLYGFSVPNSAVKGDTPKTI 240

280 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYI 339

241 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYI 300

340 DPNVKKTGNTLATIGSTTANKTVLVYKYGKGFYNSIKGTIDQIDKTNNTVQTIY 399

301 DPNVKKTGNTLATIGSTTANKTVLVYKYGKGFYNSIKGTIDQIDKTNNTVQTIY 360

400 VNPSGNVTAPVLVTGNLKNPTSDNALIDQQTISIKYKVDNAADLSESYFVNPENFEDVT 459

361 VNPSGNVTAPVLVTGNLKNPTSDNALIDQQTISIKYKVDNAADLSESYFVNPENFEDVT 420

QY 460 NSVNIFFPNQYKVEFNTPDDOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIWR 519

DB 421 NSVNIFFPNQYKVEFNTPDDOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIWR 480

QY 520 SMSWDNEVAFNNGSGSGDIDKPVVPEQDPEGEIPIEPIE 559

DB 481 SMSWDNEVAFNNGSGSGDIDKPVVPEQDPEGEIPIEPIE 520

RESULT 10

AAU75490

ID AAU75490 standard; protein; 496 AA.

XX

XX AAU75490;

XX

DT 23-APR-2002 (first entry)

DE

DE S. aureus antigenic protein associated protein #10.

XX

XX Antigenic protein; vaccine; SEREX; antibacterial; antiinflammatory;

XX dermatological; antiulcer; tuberculostatic; immunosuppressive;

XX septicaemia; food poisoning; skin disorders; peritonitis; endocarditis;

XX tuberculosis; blood infection; sepsis; meningitis; pneumonia;

XX stomach ulcer; gonorrhea; necrotising fasciitis; impetigo;

XX Lyme's disease; gastro-enteritis; dysentery; shigellosis.

XX

OS Staphylococcus aureus.

XX

XX WO200198499-A1.

XX

PD 27-DEC-2001.

XX

XX 20-JUN-2001; 2001WO-GB002685.

PF

XX 20-JUN-2000; 2000GB-00014907.

PR

XX (UYSH-) UNIV SHEPPFIELD.

PA (BIOS-) BIOSYNEXUS INC.

XX

XX Foster S, McDowell P, Brummell K, Clarke S;

PI WPI; 2002-106544-14.

DR

DR Identifying antigenic polypeptides expressed by pathogenic organisms

PT e.g., Staphylococcus aureus during infection, by SEREX (serological

PT identification of antigens by recombinant expression cloning) techniques.

XX

PS Disclosure; Page 72-74; 85pp; English.

XX

XX The invention relates to a method for identifying antigenic polypeptides

CC expressed by pathogenic organisms e.g., Staphylococcus aureus during

CC infection, by SEREX (serological identification of antigens by

CC recombinant expression cloning) techniques. The method involves providing

CC a nucleic acid library encoding genes/partial gene sequences of

CC pathogenic organisms, transforming/transfecting the library into host

CC cells, contacting the polypeptides expressed by the genes with autologous

CC antisera derived from an animal infected with, or has been infected with

CC the pathogen and purifying the nucleic acid encoding the polypeptide or

CC partial polypeptide binding to the antisera. Also included are the

CC nucleic acids and polypeptides isolated by the method, vectors and

CC transformed cells expressing them, a vaccine comprising the polypeptide

CC and the production of monoclonal antibodies against the polypeptides. The

CC protein and vaccine are useful for immunising an animal (preferably

CC human) against a pathogenic microbe. The proteins and antibodies are

CC useful for manufacturing a medicament for treating Staphylococcus aureus-

CC associated septicaemia, food poisoning or skin disorders or

CC Staphylococcus epidermidis-associated septicaemia, sepsis, meningitis,

CC endocarditis, tuberculosis, blood infections, necrotising fasciitis, impetigo,

CC pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, The present

CC sequence is an S. aureus protein sequence associated with the antigenic

CC proteins of the invention. Note: The present sequence is included in the

```
CC sequence listing but is not mentioned anywhere else in the specification
XX
SQ Sequence 496 AA;

Query Match          53.2%; Score 2514; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.9e-120;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNMKKEKHAIKRSIGVASVLVGLIGFLLSSKEADASENSVTQSDASNESKNDSS 60
DB 1 MNMKKEKHAIKRSIGVASVLVGLIGFLLSSKEADASENSVTQSDASNESKNDSS 60
QY 61 SVSAAPKTDITNVSDTKTSSNTNNGETSVAQNPAAQETQSSSTNATTEETPVTEATT 120
DB 61 SVSAAPKTDITNVSDTKTSSNTNNGETSVAQNPAAQETQSSSTNATTEETPVTEATT 120
QY 121 TTNOANTPATTTQSSNTNABELVNQTSNETTFNDNTVSSVNSPQNSTNAENVSTQDST 180
DB 121 TTNOANTPATTTQSSNTNABELVNQTSNETTFNDNTVSSVNSPQNSTNAENVSTQDST 180
QY 181 EATPSNNESAPQSDTASNDKDVVQAQVNTSAPRMAFSLAAVAADAPAAAGTDITNQLTNVT 240
DB 181 EATPSNNESAPQSDTASNDKDVVQAQVNTSAPRMAFSLAAVAADAPAAAGTDITNQLTNVT 240
QY 241 VGIDSGTIVYPHOAGYVKLYGFSVPNSAVKGDFTFKITVPKELNGLNGVSTAKVPPIMAG 300
DB 241 VGIDSGTIVYPHOAGYVKLYGFSVPNSAVKGDFTFKITVPKELNGLNGVSTAKVPPIMAG 300
QY 301 DOVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLAGTIGSTT 360
DB 301 DOVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLAGTIGSTT 360
QY 361 ANKTVLVDYKGYKFNLSIKGFIQDKTNNTYRQTIYVNPSSGDNVIAPLVTLGNLKNPT 420
DB 361 ANKTVLVDYKGYKFNLSIKGFIQDKTNNTYRQTIYVNPSSGDNVIAPLVTLGNLKNPT 420
QY 421 DSNALLDQNTSTKVKVQVNAADLSYSFVNPENFEDVNSVNIITPNPNQYKVEFNTDP 480
DB 421 DSNALLDQNTSTKVKVQVNAADLSYSFVNPENFEDVNSVNIITPNPNQYKVEFNTDP 480
QY 481 DQITTPYIVVNGHID 496
DB 481 DQITTPYIVVNGHID 496

RESULT 11
ADA89664
ID ADA89664 standard; protein; 496 AA.
AC
AC ADA89664;
XX
DT 20-NOV-2003 (first entry)
XX
DE Staphylococcus aureus antigenic protein #203.
XX
KW antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
KW antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
KW antitumor; immunostimulant; ophthalmological; pathogenic microbe;
KW bacteremia; septic shock; organ infection; skin infection;
KW bacterial basal colonisation; bacterial eye infection; septicaemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.
XX
OS Staphylococcus aureus.
XX
XX WO2003011899-A2.
XX
XX 13-FEB-2003.
XX
XX 02-AUG-2002; 2002WO-GB003606.
XX
```

```
PR 02-AUG-2001; 2001GB-00018925.
PR 09-JAN-2002; 2002GB-00000349.
XX
PA (UYSH-) UNIV SHEFFIELD.
PA (BIOS-) BIOSYNEXUS INC.
XX
PI Foster S, Mond J, Clarke S, McDowell P, Brummel K;
XX
XX WPI; 2003-256434/25.
XX
PT New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
PT useful as a vaccine for immunizing humans against e.g. bacteremia, septic
PT shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
PT impetigo.
XX
PS Claim 4; Page 161; 189pp; English.
XX
CC The present invention describes an antigenic protein or its part, which
CC is for use as a vaccine. The antigenic protein is encoded by an isolated
CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
CC sequences (designated dnaSA and dna SE, respectively; and which encodes a
CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
CC composition comprising at least one antigenic protein; (2) a method of
CC immunising an animal against a disease or condition caused by a
CC pathogenic microbe by administering the antigenic protein or the vaccine;
CC (3) an antibody or its binding part obtainable by the method above; (4)
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
CC hybridoma cell line produced by the method of (4); and (6) identifying
CC opsonic antigens expressed by a pathogenic microbe. The antigenic
CC proteins have antibacterial, neuroprotective, immunosuppressive,
CC antiinflammatory, antitumor, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteraemia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
CC bacteria-associated food poisoning, blood infections, peritonitis,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
CC strep throat, streptococcal-associated toxic shock, necrotising
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
CC skin disorders, S. epidermidis-associated septicaemia, peritonitis or
CC endocarditis. The present sequence represents a S. aureus antigenic
CC protein sequence from the present invention.
XX
SQ Sequence 496 AA;
```

```
Query Match          53.2%; Score 2514; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.9e-120;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNMKKEKHAIKRSIGVASVLVGLIGFLLSSKEADASENSVTQSDASNESKNDSS 60
DB 1 MNMKKEKHAIKRSIGVASVLVGLIGFLLSSKEADASENSVTQSDASNESKNDSS 60
QY 61 SVSAAPKTDITNVSDTKTSSNTNNGETSVAQNPAAQETQSSSTNATTEETPVTEATT 120
DB 61 SVSAAPKTDITNVSDTKTSSNTNNGETSVAQNPAAQETQSSSTNATTEETPVTEATT 120
QY 121 TTNOANTPATTTQSSNTNABELVNQTSNETTFNDNTVSSVNSPQNSTNAENVSTQDST 180
DB 121 TTNOANTPATTTQSSNTNABELVNQTSNETTFNDNTVSSVNSPQNSTNAENVSTQDST 180
QY 181 EATPSNNESAPQSDTASNDKDVVQAQVNTSAPRMAFSLAAVAADAPAAAGTDITNQLTNVT 240
DB 181 EATPSNNESAPQSDTASNDKDVVQAQVNTSAPRMAFSLAAVAADAPAAAGTDITNQLTNVT 240
QY 241 VGIDSGTIVYPHOAGYVKLYGFSVPNSAVKGDFTFKITVPKELNGLNGVSTAKVPPIMAG 300
DB 241 VGIDSGTIVYPHOAGYVKLYGFSVPNSAVKGDFTFKITVPKELNGLNGVSTAKVPPIMAG 300
QY 301 DOVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLAGTIGSTT 360
DB 301 DOVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLAGTIGSTT 360
```

Db 301 DQVLANGVIDSGNVIYTFDYNTKDDYKATLTPAYIDPENVKKTGNVLTATGIGSTT 360
QY 361 ANKTVLVDEYKGYKFNLSIKGTHIDQIDKNTNRYQTIYVNSGDNVIAPIVLTGNLKPN 420
Db 361 ANKTVLVDEYKGYKFNLSIKGTHIDQIDKNTNRYQTIYVNSGDNVIAPIVLTGNLKPN 420
QY 421 DSNALIDQNTSIKYKVDNAADLSSEYFVNPFEDVTNSVNIITFPNPNQYKVEFTPD 480
Db 421 DSNALIDQNTSIKYKVDNAADLSSEYFVNPFEDVTNSVNIITFPNPNQYKVEFTPD 480
QY 481 DQITTPYIVVNGHID 496
Db 481 DQITTPYIVVNGHID 496

RESULT 12

AD89663
ID AD89663 standard; protein; 496 AA.
XX
AC AD89663;
XX
DT 20-NOV-2003 (first entry)
XX
DE Staphylococcus aureus antigenic protein #202.
XX
KW antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
KW antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
KW antiulcer; immunostimulant; ophthalmological; pathogenic microbe;
KW bacteremia; septic shock; organ infection; skin infection;
KW bacterial basal colonisation; bacterial eye infection; septicemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.
XX

Staphylococcus aureus.

XX WO2003011899-A2.
XX 13-FEB-2003.
XX 02-AUG-2002; 2002WO-GB003606.
XX 02-AUG-2001; 2001GB-00018825.
XX 09-JAN-2002; 2002GB-00000349.
XX (NYSH-) UNIV SHEFFIELD.
XX (BQOS-) BIOSYNEXUS INC.
XX Foster S, Mond J, Clarke S, McDowell P, Brummel K;
XX WPI; 2003-256434/25.
XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
XX useful as a vaccine for immunising humans against e.g. bacteremia, septic
XX shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
XX impetigo.
XX Claim 4; Page 160-161; 189pp; English.

XX The present invention describes an antigenic protein or its part, which
XX is for use as a vaccine. The antigenic protein is encoded by an isolated
XX DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
XX sequences (designated dnaSA and dna SE, respectively; and which encodes a
XX protein expressed by a pathogenic organism. Also described: (1) a vaccine
XX composition comprising at least one antigenic protein; (2) a method of
XX immunising an animal against a disease or condition caused by a
XX pathogenic microbe by administering the antigenic protein or the vaccine;
XX (3) an antibody or its binding part obtainable by the method above; (4)
XX preparing a hybridoma cell line producing monoclonal antibodies; (5) a
XX hybridoma cell line produced by the method of (4); and (6) identifying
XX opsonic antigens expressed by a pathogenic microbe. The antigenic
XX proteins have antibacterial, neuroprotective, immunosuppressive,

CC antiinflammatory, antiulcer, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteremia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicemia, peritonitis,
CC bacteremia-associated food poisoning, blood infections, ulcers, gonorrhoea,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcer, necrotising
CC strep throat, streptococcal-associated toxic shock, gastro-enteritis,
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, S. aureus-associated septicemia, food-poisoning,
CC skin disorders, S. epidermidis-associated septicemia, peritonitis or
CC endocarditis. The present sequence represents a S. aureus antigenic
CC protein sequence from the present invention.
XX
SQ Sequence 496 AA;

Query Match 53.2%; Score 2514; DB 6; Length 496;

Best Local Similarity 100.0%; Pred. No. 1.9e-120; Indels 0; Gaps 0;
Matches 496; Conservative 0; Mismatches 0;
QY 1 MNMKKEKHAIRKKSIGVASVLGTLIGFLLSKSEADASENSVTQSDSASNESKNDSS 60
Db 1 MNMKKEKHAIRKKSIGVASVLGTLIGFLLSKSEADASENSVTQSDSASNESKNDSS 60
QY 61 SVSAAPKTDDTNSDTKTSNTNNGETSAQNPAQQTTSSTSSNTATTEPTVTEATT 120
Db 61 SVSAAPKTDDTNSDTKTSNTNNGETSAQNPAQQTTSSTSSNTATTEPTVTEATT 120
QY 121 TTQANTPATQSSNTNABELVQTSNETTENDTNTVSSVNSPONSNAENVSTQDTST 180
Db 121 TTQANTPATQSSNTNABELVQTSNETTENDTNTVSSVNSPONSNAENVSTQDTST 180
QY 181 EATPSNNEAPQSDASNKDVNQAVNTSAPRMAFSLAAVAADAPAAAGTDTITNLTNVT 240
Db 181 EATPSNNEAPQSDASNKDVNQAVNTSAPRMAFSLAAVAADAPAAAGTDTITNLTNVT 240
QY 241 VGIDSGTIVYVPHQAGYVKNLNGYFVSPNSAVKGDFTFKITVPKELNNGVTSTAKVPPIMAG 300
Db 241 VGIDSGTIVYVPHQAGYVKNLNGYFVSPNSAVKGDFTFKITVPKELNNGVTSTAKVPPIMAG 300
QY 301 DQVLANGVIDSGNVIYTFDYNTKDDYKATLTPAYIDPENVKKTGNVLTATGIGSTT 360
Db 301 DQVLANGVIDSGNVIYTFDYNTKDDYKATLTPAYIDPENVKKTGNVLTATGIGSTT 360
QY 361 ANKTVLVDEYKGYKFNLSIKGTHIDQIDKNTNRYQTIYVNSGDNVIAPIVLTGNLKPN 420
Db 361 ANKTVLVDEYKGYKFNLSIKGTHIDQIDKNTNRYQTIYVNSGDNVIAPIVLTGNLKPN 420
QY 421 DSNALIDQNTSIKYKVDNAADLSSEYFVNPFEDVTNSVNIITFPNPNQYKVEFTPD 480
Db 421 DSNALIDQNTSIKYKVDNAADLSSEYFVNPFEDVTNSVNIITFPNPNQYKVEFTPD 480
QY 481 DQITTPYIVVNGHID 496
Db 481 DQITTPYIVVNGHID 496

RESULT 13

ABJ18917
ID ABJ18917 standard; protein; 913 AA.
XX
AC ABJ18917;
XX
DT 06-MAR-2003 (first entry)
XX Pathogen specific antigen related staphylococcal protein SEQ ID No 63.
XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
XX hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
XX auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
XX autoimmune disease; HIV; hepatitis.
XX

RESULT 15	
AAE29263	
ID	AAE29263 standard; protein; 331 AA.
XX	
XX	
AC	AAE29263;
XX	
DT	27-JAN-2003 (first entry)
XX	
DE	Staphylococcus aureus Clf33 protein.
XX	
KW	Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein;
KW	immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
KW	septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.
XX	
OS	Staphylococcus aureus.
XX	
PN	WO200272600-A2.
XX	
PD	19-SEP-2002.
XX	
PF	28-JAN-2002; 2002WO-US002296.
XX	
PR	26-JAN-2001; 2001US-0264072P.

RESULT 15	
AAE29263	
ID	AAE29263 standard; protein; 331 AA.
XX	
XX	
AC	AAE29263;
XX	
DT	27-JAN-2003 (first entry)
XX	
DE	Staphylococcus aureus Clf33 protein
XX	
KW	Clumping factor A; ClfA; fibrinogen
KW	immunological; staphylococcal infect
KW	septic arthritis; sepsis; endocard
XX	
OS	Staphylococcus aureus.
XX	
PN	WO200272600-A2.
XX	
PD	19-SEP-2002.
XX	
PF	28-JAN-2002; 2002WO-US002296.
XX	
PR	26-JAN-2001; 2001US-0264072P.

26-NOV-1997; 97US-0066815P.
31-AUG-1998; 98US-0098427P.
(INHI-) INHIBITEX INC.
(ORF-) FORFAS T/A BIORESEARCH IRELAND.
(TEXA) UNIV TEXAS A & M.
(PATT) PATTI J M.
(FOST) FOSTER T J.
(JOSE) JOSEFSSON E.
(EIDH) EIDHIN D N.
(HOOK) HOOK M A O.
(PERK) PERKINS S E.

Patti JM, Foster TJ, Josefsson E, Eidhin DN, Hook MAO;
Perkins SE;
WPI; 1999-357844/30
N-PSDB; AAX77591.

Staphylococcus aureus fibrinogen-binding proteins for treating
septicemia, osteomyelitis, mastitis or endocarditis.

Claim 2; Fig 5; 143pp; English.

This invention describes novel Staphylococcus aureus fibrinogen-binding
proteins that bind both the alpha and beta fibrinogen chains. The
proteins (and their encoding nucleic acids are ClfB, SdrC, SdrD and
SdrE). Staphylococcus aureus is thought to utilize fibrinogen to adhere
to medical devices, binding proteins that bind both the alpha and beta
fibrinogen chains (ClfB, SdrC, SdrD and SdrE). Antibodies against ClfB,
competitive inhibitors to block this binding. Antibodies against ClfB,
SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.
The proteins of the invention can be used in a pharmaceutical composition
for the treatment of Staphylococcus aureus infection e.g. septicemia,
osteomyelitis, mastitis or endocarditis or to inhibit the binding of S.
aureus to the extracellular matrix. The proteins or their fragments may
be used to coat a medical device to reduce the S. aureus infection of an
indwelling medical device, especially where the medical device is
selected from the group consisting of vascular grafts, vascular stents,
intravenous catheters, artificial heart valves, and cardiac assist
devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or
encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
region of a gene encoding it may be used as an identifying probe for the
identification of genes and encoding proteins from Staphylococcus aureus
(other than ClfA). S. hemolyticus, S. lugdenensis, and S. schlereria.
The proteins of the invention have antibacterial activity. (Updated on 20
-MAR-2003 to correct FA field.)

XX
XX
SQ

Sequence 918 AA;

Query Match 38.3%; Score 1808.5; DB 2; Length 918;
Best Local Similarity 43.5%; Pred. No. 3.8e-84;
Matches 429; Conservative 147; Mismatches 269; Indels 141; Gaps 31

QY 5 KXKHAIRKXSGVAVLVLTGLIGLLSKSEADASE--NSVTQSDASNESKNSDSSV 62
Db 16 KQKYSIRRFVTGVTTSVIGVATILFG-IGNHQAQSEQNSDITQS-SKNNASADEKKNM 73

QY 63 SAAPKTDITNVSDTKTSNTNNGETSVQAQNPQAQOBTTOSSSTN-ATTETPTVTEATT 121
Db 74 IETPQLNTANDTSDISANTSNANVSDTTKPMSTQTSNTTTEPATNETP----QPTAI 129

QY 122 TNQANTPATTOSSNTNABELVNOQTSNETTFNDNTVSSVNSPQNSTNAENVSTQDTSTE 181
Db 130 KQQAT--AAKMQDQTVPOEGNSQVDNKTT-NDANSI-----ATNSE----LKNSQTL 174

QY 182 ATPSNNSAPOSTDASNKDQVNVQAVNTSAPRMWAFSLAAVAADAP-----AAGTDITNQ 235
Db 175 DLP---QSSPQT-----ISNAOQTSPKSVRTAVRSLVAEAPVNVNADARGNTVNDK 223

QY 236 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
Db 237 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 238 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 239 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 240 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 241 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 242 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 243 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 244 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 245 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 246 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 247 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 248 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 249 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 250 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 251 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 252 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 253 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 254 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 255 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 256 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 257 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 258 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 259 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 260 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 261 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 262 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 263 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 264 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 265 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 266 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 267 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 268 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 269 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 270 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 271 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 272 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 273 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 274 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 275 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 276 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 277 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 278 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 279 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 280 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 281 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 282 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 283 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 284 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 285 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 286 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 287 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 288 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 289 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 290 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 291 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 292 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 293 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 294 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 295 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 296 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 297 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 298 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 299 LTNVTNVGIDSGTIVYPHAGYVVKLVNGFSVPNSAVKGDTFKITVPKELMNGVTSTAKVP 295
XX 300 L

PR 12-MAR-2001; 2001US-0274611P.
PR 18-JUN-2001; 2001US-0298413P.
PR 30-JUL-2001; 2001US-0308116P.
XX (INH1-) INHIBITEX INC.
FA Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
XX
XX WPI; 2002-759834/82.
DR N-PSDB; AAD46862.
XX
XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for
PT treating or preventing Staphylococcus aureus infection e.g. wound
PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
PT a human or animal.
XX
PS Claim 9; Page 70-72; 80pp; English.
XX
XX The invention relates to monoclonal antibody which binds the clumping
CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
CC monoclonal antibody is useful for treating or preventing S. aureus
CC infection in a human or animal, and for inhibiting the binding of
CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
CC S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3
CC protein is useful for inducing an immunological response in a human or
CC animal. These staphylococcal infections include wound infections, sepsis,
CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC present sequence is Staphylococcus aureus Clf33 protein
XX
SQ Sequence 331 AA;
Query Match 36.6%; Score 1727; DB 5; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.7e-80;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 221 VAADAPAGTDTITNLTNTVTGIDSGTIVYHQAGYVKLNYSFVSPNSAVKGDTEKITVP 280
Db 2 VAADAPAGTDTITNLTNTVTGIDSGTIVYHQAGYVKLNYSFVSPNSAVKGDTEKITVP 61
QY 281 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKALTMPAYID 340
Db 62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKALTMPAYID 121
QY 341 PENVKKTGNVTIATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYQTIYV 400
Db 122 PENVKKTGNVTIATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYQTIYV 181
QY 401 NPSGDNVIAPVLTGNLKPNTDSNALIDQNTSIVKYKVDNADLSSESFVNPFEDVTN 460
Db 182 NPSGDNVIAPVLTGNLKPNTDSNALIDQNTSIVKYKVDNADLSSESFVNPFEDVTN 241
QY 461 SVNITFPNPNQYKVFNTPDQITTPYIVVNGHIDPNSKGDALRSTLYGYSNIIWRS 520
Db 242 SVNITFPNPNQYKVFNTPDQITTPYIVVNGHIDPNSKGDALRSTLYGYSNIIWRS 301
QY 521 MSWDNEVAFNNGSGGDKPKVPEQPD 550
Db 302 MSWDNEVAFNNGSGGDKPKVPEQPD 331

Search completed: November 3, 2004, 19:02:59
Job time : 181.713 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 19:03:11 ; Search time 13.6742 Seconds
(without alignments)
1078.240 Million cell updates/sec

Title: US-09-679-643-2
Perfect score: 4725
Sequence: 1 MNMKKEHAKKISIGVAS.....LASIGSLFLRRKKENDKK 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 66696 seqs, 15802848 residues

Total number of hits satisfying chosen parameters: 66696

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:
1: /cgn2_6/ptodata/2/paa/pct/US04/02460-3
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pcp:
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pcp:
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pcp:
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pcp:
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pcp:
7: /cgn2_6/ptodata/2/paa/US11_NEW_COMB.pcp:
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pcp:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	343.5	7.3	1903	1	PCT-US04-02460-3	Sequence 3, Appli
2	313	6.6	2110	6	US-10-732-923-8666	Sequence 8666, Ap
3	297	6.3	5322	6	US-10-732-923-8729	Sequence 8729, Ap
4	275	5.8	1003	6	US-10-732-923-3319	Sequence 3319, Ap
5	270	5.7	733	6	US-10-732-923-2268	Sequence 2268, Ap
6	265.5	5.6	2165	6	US-10-732-923-13547	Sequence 13547, A
7	264.5	5.6	248	6	US-10-220-366A-21871	Sequence 21871, A
8	263	5.6	1883	6	US-10-732-923-8730	Sequence 8730, Ap
9	263	5.6	1883	6	US-10-732-923-8731	Sequence 8731, Ap
10	260	5.5	1808	6	US-10-732-923-8793	Sequence 8793, Ap
11	259.5	5.5	3328	6	US-10-732-923-8311	Sequence 8311, Ap
12	258.5	5.5	2288	6	US-10-732-923-8881	Sequence 8881, Ap
13	258.5	5.5	2337	6	US-10-732-923-8888	Sequence 8888, Ap
14	258.5	5.5	2375	6	US-10-732-923-8880	Sequence 8880, Ap
15	258.5	5.5	2454	6	US-10-732-923-8894	Sequence 8894, Ap
16	258.5	5.5	2492	6	US-10-732-923-8882	Sequence 8882, Ap
17	257.5	5.4	2492	6	US-10-732-923-8881	Sequence 8881, Ap
18	257.5	5.4	2492	6	US-10-732-923-8883	Sequence 8883, Ap
19	255	5.4	1081	6	US-10-732-923-8080	Sequence 8080, Ap
20	253.5	5.4	974	1	PCT-US04-02460-2	Sequence 2, Appli
21	253.5	5.4	1709	6	US-10-732-923-8828	Sequence 8828, Ap
22	251.5	5.3	1806	6	US-10-732-923-8671	Sequence 8671, Ap
23	250.5	5.3	2476	6	US-10-732-923-8136	Sequence 8136, Ap
24	249.5	5.3	2307	6	US-10-732-923-19147	Sequence 19147, A
25	249.5	5.3	2314	6	US-10-732-923-19146	Sequence 19146, A

ALIGNMENTS

RESULT 1

PCT-US04-02460-3
; Sequence 3, Application PC/TUS0402460

; GENERAL INFORMATION:

; APPLICANT: Chang, Chia-Hwa

; APPLICANT: Liu, Xiaowen

; APPLICANT: Lewicki, John A.

; APPLICANT: Xu, Qiang

; APPLICANT: Osel, Inc.

; TITLE OF INVENTION: Surface Expression of Biologically Active Proteins in

; FILE OF INVENTION: Bacteria

; FILE REFERENCE: 016976-000810PC

; CURRENT APPLICATION NUMBER: PCT/US04/02460

; CURRENT FILING DATE: 2004-01-28

; PRIOR APPLICATION NUMBER: US 60/443,619

; PRIOR FILING DATE: 2003-01-29

; NUMBER OF SEQ ID NOS: 89

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 1903

; TYPE: PRT

; ORGANISM: Lactobacillus jensenii 1153

; FEATURE:

; OTHER INFORMATION: genomic C370 cell wall anchor sequence

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (1309)..(1903)

; OTHER INFORMATION: CWA200 cell wall associated region with anchor

; OTHER INFORMATION: motif

; FEATURE:

; NAME/KEY: PEPTIDE

; LOCATION: (1869)..(1872)

; OTHER INFORMATION: anchor motif, sorting signal, cell wall targeting

; OTHER INFORMATION: region

PCT-US04-02460-3

Query Match 7.3%; Score 343.5; DB 1; Length 1903;

Best Local Similarity 18.7%; Pred. No. 1.6e-11;

Matches 190; Conservative 181; Mismatches 454; Indels 189; Gaps 29;

QY 33 SSKEADAGENSVTQSDASNESKNDSS-----SVSAAPKTDNTVSDTTSNTNNG 85

Db 964 SAKDAISANTVDEITTAKNNAVSTDTNNMKLSEDTKLAADKTQDPYLNADLKQAYDK 1023

QY 86 ETSVQANPAQQTQSSSTNATETPTVGEATTTTNCANTPATQSSNTNAAELVNGT 145

Db 1024 AVEEAQKLLNKETGTSGADKDPAEVARIKQAVDDAYDALNGNSLDDAKQAKQAVDK- 1082

```
QY 146 SNETTEND-----TNTVSSVNSPQNSTNAE--NVSTTOD-TSTETATP 184
Db 1083 -NYINLNDKQKETAKKRIDSAKSDEYVNNADKINSGLNEXKMGELKEVSNLSDKIETT-SNY 1141
QY 185 SNNES-----APQSTASNKDVVNOAVNTSAPRMAFSAVAADAPA 227
Db 1142 SNADSDKKQAYKETADKIHTETAPSGDDLTTDDVNNLITDEATKR-----AALNGDARE 1195
QY 228 AGTDITNQLTNVTVIDSGTTV---YPHQAGYVKLYNGFVSPNSAVKGDFTFKITVPKELN 284
Db 1196 KARQELNNYNSKSLQGGSTLDPRIYNASEEKKQAFQKALDNKAKALDNETTEAAYKS 1255
QY 285 LNGVTSTAKVPPIMAGQVLANGVID---SDGNVIYFTFDYVNTKDDVKATLTPAYIDP 341
Db 1256 ANDELQAKAD---LDCQTTTKSKLDIAIKDANNAKGTDKYNASDDTKSKFDALKA 1311
QY 342 ENUKTGNVTLATIGSTTANKTVLVYKEYKGFYNNLSIKGTIDQIDKTNNTVROTIVYN 401
Db 1312 EEVKNNNSNATQ-----KEVDDATNNLKQA----- 1335
QY 402 PSGDNVIAPLVTGNLKPTDTSNALID---QONTSIKVKYVDNAADLSESYFYVNP-ENFE 456
Db 1336 -----QNNLNGQTTDKSKLDIAIKDANNAKGTDKYNASDDTKSKFDALKA 1384
QY 457 DVINSVNITPPN---PNQYKVFENTPDDQ-----ITPYIVVNGH 494
Db 1385 EVKNNNSNATQKEVDDATNNLKQAQNLDLGGTTDKSKLDEAITDANNKLTDKY---NNAS 1441
QY 495 IDPNSKGDALRSTLYGYSNIIWRSMWNEV--APNNGSGGDDGIDKVPVPEQDDEPG 552
Db 1442 DTKSKFDEALKKA-----ENVKNSNATQKEVDDATNNLKQAQNLDL-----GQTTDKS 1491
QY 553 EIEPIPEDSDSPGSGSDNSDSGS-----DSGSDSTSDSGSDSASD 600
Db 1492 KLDEAITDANNKSTDKYNNASDDTKSKFDEALKKAEVKNNSNATQKEVDDATNNLKQA 1551
QY 601 SDSDSASDSASDSASDSASDSDNPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 660
Db 1552 QNNLDGQTTDKSKLDEAITDANNKSTDKYNNASDDTKSKFD--DALKKAEEVKNNNSNATQ 1610
QY 661 SDSDSDSDSDSDSDSDSDS--SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 714
Db 1611 KEVDDATNNLKQAQNLDLGGTTNKTDLNDAIKDANAKGTDKYKNASDDTKSKLDETLKK 1670
QY 715 ----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 770
Db 1671 ABEVKNNNSNATQKEVDDATNNLKQAQNLDLGGTTDKSKLDEAIKSDADTKSTDKYNNASD 1730
QY 771 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSSES---SDSDSDSDSDSDS 827
Db 1731 DTKSKFDEALKKAEVKNNSNATQKEVDDATNNLKQAQNLDLGGTTNKTDLNDAIKDANNA 1790
QY 828 SDSDSDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 878
Db 1791 AKGTDKYNNSDDTKSKFDALKKAEEVKNNNSNATQKEVDDATNNLKNTLNNLKGPAPKK 1850
QY 879 NGTNASKNNEAKDSKEPLDPTGSEDEANTSLTWGLIASIGSLLLPRKKENKOK 932
Db 1851 ANLIASKONAKIHKQTLPLQTGTETNPLTAIGLGLMA-LGA-GIFAKKKRKKDE 1902
```

RESULT 2

```
US-10-732-923-8666
; Sequence 8666, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
```

```
; SEQ ID NO 8666
; LENGTH: 2110
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-732-923-8666
```

```
Query Match 6.6%; Score 313; DB 6; Length 2110;
Best Local Similarity 18.3%; Pred. No. 8.1e-10;
Matches 101; Conservative 117; Mismatches 192; Indels 142; Gaps 15;
```

```
QY 360 TANKTVLVDYKEYKGFYNNLSIKGTIDQID--KUNNTYR--QTIYVNPNGDNVIAPLVTGN 415
Db 166 TLNKKKEMRFKPF-----VKEAIKKIKDARHNAHELFPENKYSVSSNN----- 209
QY 416 LKNTSINALIDQONTISI-----KVYKVDNAADLSESYFYVNPENPFDVNTVITPPN 471
Db 210 -----NSETVNNASSVDNGDKELKBDLTDQBEDYLLDEQMSSTDESEN----- 255
QY 472 YKVEFNTPDDQITTPYIVVNGHIDPNSKGDALRSTLYGYSNIIWRSMWNEVAFNN 531
Db 256 KEKEINMLDDEANLPI-----EELKRMYGFKS----- 283
QY 532 GSGSGPIDKPVVPEQDEPGEIEPIPEDSDSPGSDSDSDSDSDSDSDSDSDSDSDSDSGS 591
Db 284 -----GEDYINFWENEDDANE-ENVLETSHNDEKSGDNSIGEDDNNND-----EKGG 329
QY 592 DSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSAS 651
Db 330 DNNIDEDDNNNDKSGNSIGEDDNNNDHKSNGDNNIDED--DNNNDHKSSEDNSIGEDDNN 388
QY 652 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 698
Db 389 DEKGGNNIDENNNDHKSSEDNNIDENNNDHKSSEDNNIDENNNDHKSSEDNNIDENNND 448
QY 699 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 752
Db 449 KNYNKTGTDYKSDENYGPTRFHNKLKKEKYDEYDTKLKIEKREENKNYKDEHEYES 508
QY 753 SDSDS-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 791
Db 509 DNYDKKINKKKELIILLKNDIENDSDSETSEHKRDSRSCQKQNCCKKRII1KDEYNLR 568
QY 782 -----DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 834
Db 569 TKIAKSPSDNNNSDNNNDNNNDNNNDNNNDNNNDNNNDNNNDNNNDNNNDNNNDNNNDNN 628
QY 835 SDSASDSDSGD 846
Db 629 NEHKNSDDND 640
```

RESULT 3

```
US-10-732-923-8729
; Sequence 8729, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8729
; LENGTH: 5322
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-732-923-8729
```

```
Query Match 6.3%; Score 297; DB 6; Length 5322;
Best Local Similarity 18.6%; Pred. No. 1.7e-08;
Matches 217; Conservative 165; Mismatches 461; Indels 322; Gaps 41;
```

[illegible][illegible]

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:54:26 ; Search time 45.7559 Seconds
(without alignments)
1352.277 Million cell updates/sec

Title: US-09-679-643-2
Perfect score: 4725
Sequence: 1 MNMKKEKHAIRKKSIGVAS.....LASIGSLFLFRKKNKK 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4725	100.0	933	3	US-08-293-728-2
2	4725	100.0	933	3	US-09-421-868-2
3	4662	98.7	936	4	US-08-956-171E-5249
4	4662	98.7	936	4	US-08-781-986A-5249
5	1810.5	38.3	918	4	US-09-200-850E-1
6	1727	36.6	345	3	US-08-856-253-7
7	1628	34.5	1092	4	US-09-147-405B-15
8	1584.5	33.5	1742	4	US-09-386-962C-4
9	1584.5	33.5	1742	4	US-09-386-959-4
10	1414	29.9	1166	4	US-09-200-650E-7
11	1241.5	26.3	930	4	US-09-200-650E-3
12	1110.5	23.5	1315	4	US-09-200-650E-5
13	880	18.6	930	3	US-09-134-001C-5314
14	880	18.6	930	4	US-09-386-962C-10
15	847.5	17.9	2137	3	US-09-134-001C-4463
16	799	16.9	287	4	US-09-710-279-468
17	695	14.7	486	4	US-09-710-279-788
18	671	14.2	487	4	US-09-386-962C-14
19	671	14.2	487	4	US-09-386-959-65
20	614	13.0	1027	4	US-08-956-171E-5254
21	614	13.0	1027	4	US-08-781-986A-5254
22	530.5	11.2	1155	4	US-09-710-279-1780
23	488	10.3	466	3	US-09-134-001C-4749
24	472.5	10.0	669	4	US-09-107-532A-6532
25	452	9.6	582	4	US-09-147-405B-13
26	452	9.6	593	4	US-09-147-405B-11
27	413	8.7	414	4	US-09-248-796A-19046

28	388.5	8.2	556	4	US-09-248-796A-22338	Sequence 22338, A
29	379	8.0	2870	4	US-09-479-467A-15	Sequence 15, Appl
30	379	8.0	3178	4	US-09-479-467A-4	Sequence 4, Appl
31	350	7.4	461	4	US-09-248-796A-23039	Sequence 23039, A
32	346	7.3	827	4	US-09-248-796A-17307	Sequence 17307, A
33	344.5	7.3	1060	3	US-08-911-393-2	Sequence 2, Appl
34	344.5	7.3	1060	4	US-09-955-909-2	Sequence 44, Appl
35	343.5	7.3	447	4	US-09-794-422-44	Sequence 2, Appl
36	334.5	7.1	886	4	US-08-956-171E-5235	Sequence 44, Appl
37	334.5	7.1	886	4	US-08-781-986A-5235	Sequence 5235, Ap
38	319.5	6.8	699	4	US-09-248-796A-23200	Sequence 5235, Ap
39	302.5	6.4	917	4	US-09-248-796A-18960	Sequence 23200, A
40	300	6.3	1776	4	US-09-556-877-179	Sequence 18960, A
41	300	6.3	1776	4	US-09-620-412C-179	Sequence 179, App
42	300	6.3	1776	4	US-09-598-419-179	Sequence 179, App
43	299.5	6.3	821	4	US-09-556-877-195	Sequence 195, App
44	299.5	6.3	821	4	US-09-620-412C-195	Sequence 195, App
45	299.5	6.3	821	4	US-09-598-419-195	Sequence 195, App

ALIGNMENTS

RESULT 1
US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match	100.0%;	Score	4725;	DB 3;	Length	933;			
Best Local Similarity	100.0%;	Pred. No.	1.9e-293;						
Matches	933;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MNMKKEKHAIRKKSIGVASVLVTGLIGFLLSKEADASENSVTQSDSASNESKNDSS	60						
Db	1	MNMKKEKHAIRKKSIGVASVLVTGLIGFLLSKEADASENSVTQSDSASNESKNDSS	60						
Qy	61	SVGAAPKTDITNYSDDTKTSNTNNGETSVAQNPAQOETTSQSSSTNATTEETPTVGEATTT	120						
Db	61	SVGAAPKTDITNYSDDTKTSNTNNGETSVAQNPAQOETTSQSSSTNATTEETPTVGEATTT	120						
Qy	121	TTNQANTPATTSQSSNTNAEELVNQTSNETTFNDNTVSSVNSPONSNAENVTQDTSST	180						
Db	121	TTNQANTPATTSQSSNTNAEELVNQTSNETTFNDNTVSSVNSPONSNAENVTQDTSST	180						
Qy	181	EATPSNNEAPQSDTASNKDVQNAVNTSAPRMEAFSLAAVAADAPAAAGDITNQLTNVT	240						
Db	181	EATPSNNEAPQSDTASNKDVQNAVNTSAPRMEAFSLAAVAADAPAAAGDITNQLTNVT	240						
Qy	241	VGDSGTTVPHQAGYVKNLNGFVSPNSAVKGTFTKTPKELNNGVTSTAKVPPIMAG	300						
Db	241	VGDSGTTVPHQAGYVKNLNGFVSPNSAVKGTFTKTPKELNNGVTSTAKVPPIMAG	300						
Qy	301	DQVLANGVIDSGNVIYFTFDYNTKDDVKATLTPPAYIDPENVKKTGNVLTATGIGSTT	360						
Db	301	DQVLANGVIDSGNVIYFTFDYNTKDDVKATLTPPAYIDPENVKKTGNVLTATGIGSTT	360						
Qy	361	ANKTVLVDYKYGKFNLSIKGTIDQDKNTNNTYRQIYINPNSGDNVIAPVLGTGNLKENT	420						
Db	361	ANKTVLVDYKYGKFNLSIKGTIDQDKNTNNTYRQIYINPNSGDNVIAPVLGTGNLKENT	420						

appl, cat

COUNTRY: USA
ZIP: 20850
COMPUTER/READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICANT: Charles Kunsch
FILING DATE: January 5, 1996
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5249:

US-08-956-171E-5249
Query Match 98.7%; Score 4662; DB 4; Length 936;
Best Local Similarity 99.0%; Pred. No. 2e-289;
Matches 924; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY	1	MNNKKKEKHAIKKSIGVASVLGTLIGFLLGSKSEADASENSVTQSDASNESKNDSS	60
Db	10	MNNKKKEKHAIKKSIGVASVLGTLIGFLLGSKSEADASENSVTQSDASNESKNDSS	69
QY	61	SVSAAPKTDITDVTSDTKTSNTNNGETSAQNPAQQTTSQSSNTATTEPTVGEATT	120
Db	70	SVSAAPKTDITDVTSDTKTSNTNNGETSAQNPAQQTTSQSSNTATTEPTVGEATT	129
QY	121	TINQANTPATQSSNTNABELVNQTSNETTSDNTNTVSSVNSPQNSTNAENVSTTQD	180
Db	130	TINQANTPATQSSNTNABELVNQTSNETTSDNTNTVSSVNSPQNSTNAENVSTTQD	189
QY	181	EATPSNNESAPQSDTASNDKDVVQAVNTSAPRMAFSLAAVAADAPAAAGTDITNQL	240
Db	190	EATPSNNESAPQSDTASNDKDVVQAVNTSAPRMAFSLAAVAADAPAAAGTDITNQL	249
QY	241	VGIDSGTIVPHQAGYVKLYGFSVPNSAVKGDFTKITVPKELNGLNGVTSTAKVPP	300
Db	250	VGIDSGTIVPHQAGYVKLYGFSVPNSAVKGDFTKITVPKELNGLNGVTSTAKVPP	309
QY	301	QVILANGVIDSDGNVIITFTDYNTKDDVKATLTMPAYIDPENVKKTGNVLTGIGST	360
Db	310	QVILANGVIDSDGNVIITFTDYNTKDDVKATLTMPAYIDPENVKKTGNVLTGIGST	369
QY	361	ANKTLVDYKYGKFNLSIKGTIDQDKTNTYRTIYVNPSPGDNVIAVLTGNLKPNT	420
Db	370	ANKTLVDYKYGKFNLSIKGTIDQDKTNTYRTIYVNPSPGDNVIAVLTGNLKPNT	429
QY	421	DSNALIDQQNTSIKVKYVDNAADLSSEYFVNPNFEDVTNSVNTTFPNQYKVFNTPD	480
Db	430	DSNALIDQQNTSIKVKYVDNAADLSSEYFVNPNFEDVTNSVNTTFPNQYKVFNTPD	489
QY	481	DQITPTPIVVVNGHIDPNSKGLALRSTLYGYNSTIIWRSMWSDNEVAFNNGSGSGD	540
Db	490	DQITPTPIVVVNGHIDPNSKGLALRSTLYGYNSTIIWRSMWSDNEVAFNNGSGSGD	549

QY	541	KPVVPEQPDPEGEIEPIPEDSDSDPGSDSGSDNSDSDSGSDSTSDSGSDASDSDSA	600
Db	550	KPVVPEQPDPEGEIEPIPEDSDSDPGSDSGSDNSDSDSGSDSTSDSGSDASDSDSA	609
QY	601	SDSDASD	660
Db	610	SDSDASD	669
QY	661	SD	720
Db	670	SD	729
QY	721	SD	780
Db	730	SD	789
QY	781	SD	840
Db	790	SD	843
QY	841	SDSGSDSDSDSDSDSDSDSDSGSDSGSDNSDSDSGSDNSDSDSGSDNSDSDSGSD	900
Db	844	SDSGSDSDSDSDSDSDSDSDSGSDSGSDNSDSDSGSDNSDSDSGSDNSDSDSGSD	903
QY	901	SEDEANTSLIWGLLASIGSILLFRKKENKDKX	933
Db	904	SEDEANTSLIWGLLASIGSILLFRKKENKDKX	936

RESULT 4
US-08-781-986A-5249
Sequence 5249, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5249

Query Match

98.7%; Score 4662; DB 4; Length 936;


```

Best Local Similarity 99.0%; Pred. No. 2e-289;
Matches 924; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 1 MNMKKEKHAIRKKSIGVASVLVGLTIGFLLSSKEADASENSVTQSDASNESKNSDSS 60
Db 10 MNMKKEKHAIRKKSIGVASVLVGLTIGFLLSSKEADASENSVTQSDASNESKNSDSS 69
QY 61 SVSAAPKTDITNVSTKTSNTNNGETSVAQNPACQETQSSSTNATTEETPVTEATT 120
Db 70 SVSAAPKTDITNVSTKTSNTNNGETSVAQNPACQETQSSSTNATTEETPVTEATT 129
QY 121 TTNQANTPATTOSSNTNABELVNOTSNETTFNDTNTVSSVNSPQNSTNAENVSTTQDST 180
Db 130 TTNQANTPATTOSSNTNABELVNOTSNETTFNDTNTVSSVNSPQNSTNAENVSTTQDST 189
QY 181 EATPSNNEAPOSTDASNDKOVNQAVNTSAPWRAPSLAAVAADAPAGTDTITNLTNVT 240
Db 190 EATPSNNEAPOSTDASNDKOVNQAVNTSAPWRAPSLAAVAADAPAGTDTITNLTNVT 249
QY 241 VGIDSGTTVYHQAGYVKNLYGFSVPNSAVKGDTEFKITVPKELNLNGVTSTAKVPPIMAG 300
Db 250 VGIDSGTTVYHQAGYVKNLYGFSVPNSAVKGDTEFKITVPKELNLNGVTSTAKVPPIMAG 309
QY 301 DOVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYIDPNVKKTGNTVLTATGIGSTT 360
Db 310 DOVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYIDPNVKKTGNTVLTATGIGSTT 369
QY 361 ANKTVLVDEYKYGKFNLSIGTIDIDKTNNTYEQTIYVNSPQNVAPVLTKNLKNT 420
Db 370 ANKTVLVDEYKYGKFNLSIGTIDIDKTNNTYEQTIYVNSPQNVAPVLTKNLKNT 429
QY 421 DSNALIDQNTSIKYKYVDNAADLSSEYFVNPENFEDVTNSVNIIFPNPNQYKVFNTPD 480
Db 430 DSNALIDQNTSIKYKYVDNAADLSSEYFVNPENFEDVTNSVNIIFPNPNQYKVFNTPD 489
QY 481 DOIITPYIVVNGHIDPNKGDALRSTLYGNSNI IWRMSWMDNEVAFNNGSGGDGID 540
Db 490 DOIITPYIVVNGHIDPNKGDALRSTLYGNSNI IWRMSWMDNEVAFNNGSGGDGID 549
QY 541 KPVPPEQDEPEGEIEPIEDSDSDSGSDSDSGSDSDSGSDSDSGSDSDSDSDSDSD 600
Db 550 KPVPPEQDEPEGEIEPIEDSDSDSGSDSDSGSDSDSGSDSDSGSDSDSDSDSDSD 609
QY 601 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 660
Db 610 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 669
QY 661 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 720
Db 670 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 729
QY 721 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 780
Db 730 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 789
QY 781 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 840
Db 790 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 843
QY 841 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Db 844 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 903
QY 901 SEDEANTSLIWGLLASIGSLIIIFRKKENKDKK 933
Db 904 SEDEANTSLIWGLLASIGSLIIIFRKKENKDKK 936

```

```

RESULT 5
US-09-200-650E-1
; Sequence 1, Application US/09200650E
; Patent No. 6680195
; GENERAL INFORMATION:

```

```

; APPLICANT: Patti, Joseph M.
; APPLICANT: Poster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Eidhinn, Deirdre Ni
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 918
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-200-650E-1

Query Match 38.3%; Score 1810.5; DB 4; Length 918;
Best Local Similarity 43.5%; Pred. No. 1.8e-107;
Matches 429; Conservative 148; Mismatches 268; Indels 141; Gaps 31;

QY 5 KKEKHAIRKKSIGVASVLVGLTIGFLLSSKEADASE--NSVTQSDASNESKNSDSSV 62
Db 16 KQNKVSIIRFTVGTTSVIVGATILFG-IGNHQAQASEQSNDDTQSQ-SKNNASADEKNNM 73
QY 63 SAAPKTDITNVSDTKTSSNTNNGETSAQNPAQOETQSSSTN-ATTEETPVTEATT 121
Db 74 IETPQLNTANDTSDISANTNSANVDSTTKPMSTQTSNTTTEPASTNETP---QPTAI 129
QY 122 TTNQANTPATTOSSNTNABELVNOTSNETTFNDTNTVSSVNSPQNSTNAENVSTTQDST 181
Db 130 KNOAT--AAKMQDQTVPOEGNSQVDNKT--NDANSI-----ATNSE---LNKSQL 174
QY 182 ATPSNNEAPOSTDASNDKOVNQAVNTSAPWRAPSLAAVAADAP-----AAGTDITNQ 235
Db 175 DLP---QSSPT-----ISNAQGTISKPSVRFVAVSLAVAEVNVAAAKGNVNDK 223
QY 236 LTNTVGTIDSGTTVYHQAGYVKNLYGFSVPNSAVKGTPTKITVPKELNLNGVTSTAKVP 295
Db 224 VTASNFLEK-TTFDPNQSGNTFMAANFTVTDKVKSGDYF-----TAKLP 267
QY 296 PIMAGDQVLANGVID-----SDGNVI-----YTFDYVNTKD 327
Db 268 ----DSLGTGSDVDYSSNNMTPIADIKSTNGDVAKATYDILTKTVTFVFTDYNNKB 322
QY 328 DVKATLTMPAYIDPNVKKTGNTVLTATGIGSTTANKTFLVDYK-----YGFYNLISKGT 383
Db 323 NINGQFSLPFTDRAPKAPKSGTYDANINIADEMFMFNKITYNYSSPIAGIDKPGNANISQ 382
QY 384 IDQIDKN--NTYQTIYVNP-----GDNVIAPVLTKNLKENTDSNALIDQNTSIKYVK 437
Db 383 IIGVDTASGQNTYQTVFVNPQKRVLGNTWV--YKGYODKIERSSGKVSATDTKLRIPE 440
QY 438 VDNAADLSSEYFVNP--ENPEDVTNSV--NITFPNPQYKVFNTPDQITPTPIYVNG 493
Db 441 VNDTSKLSDSYAADPNDSNKLKEVTDQFKRIYYEHPNVASIKFG---DITKTYVVLVEG 496
QY 494 HIDPNSKGDALRLRGLTYGN-----SNIWRMSWMDNEVAFNNGSGSDGIDKPVVPEOPD 549
Db 497 HYDNTGN--LKTQVILQENVDPVTRDYSLIFGNWNNENVYGGGADG--DSAVNPKDPT 552
QY 550 EPGIEPIPE--DSDDPGSGSGSDNSDSDSGSDSTSDSGSDSASDSDSDSASD 608
Db 553 PGPPVDPEPSPDPEPEPTD-----PEPSPDPEPSPDPPDSDSD 594
QY 609 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 668
Db 595 SDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 654

```

QY 669 SDS 728
DB 655 SDS 714
QY 729 SDS 788
DB 715 SDS 774
QY 789 SDS 848
DB 775 SDS 834
QY 849 SSSDSSESNSDSSESNNVPPNSKNGTNASNKNEAKDSKEPLPTGSEDE-ANT 907
DB 835 SDSDSDSDSR---VTPPNEQKAPSNPKGEVNHNSKVKQHKTDALPETGKSENTNA 891
QY 908 SLTWGLLASTGSLLLFRKXENKDKK 933
DB 892 TLFGAMWALLGSLLLFRKXQDHKEK 917

RESULT 6
US-08-856-253-7

; Sequence 7, Application US/08856253
; Patent No. 6288214

GENERAL INFORMATION:

; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:

NAME:

; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK-193

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO:

; 7:

SEQUENCE CHARACTERISTICS:

; LENGTH: 345 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-856-253-7

Query Match 36.6%; Score 1727; DB 3; Length 345;

Best Local Similarity 100.0%; Pred. No. 1.1e-102;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDFTKITVP 280

DB 14 VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDFTKITVP 73
QY 281 KEINLNGVISTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID 340
DB 74 KEINLNGVISTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID 133
QY 341 PENVKKTGNVTLAGTIGSTTANKTIVLDVYKYGKFNLSIKGTIDQIDKTNNTYRQTIYV 400
DB 134 PENVKKTGNVTLAGTIGSTTANKTIVLDVYKYGKFNLSIKGTIDQIDKTNNTYRQTIYV 193
QY 401 NPSGDNVIAVPLTGNLKPNNTDSNALIDQQNTSIKVKYKVDNAADLSSESFYVNPENFEDVTN 460
DB 194 NPSGDNVIAVPLTGNLKPNNTDSNALIDQQNTSIKVKYKVDNAADLSSESFYVNPENFEDVTN 253
QY 461 SVNITFPNPQYKVEFTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGNSNIWRS 520
DB 254 SVNITFPNPQYKVEFTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGNSNIWRS 313
QY 521 MSWDNEVAFNNGSGSDGIDKPVVPEQDPE 550
DB 314 MSWDNEVAFNNGSGSDGIDKPVVPEQDPE 343

RESULT 7

US-09-147-405B-15

; Sequence 15, Application US/09147405B

; Patent No. 6733758

GENERAL INFORMATION:

; APPLICANT: Guss, Bengt

; APPLICANT: Nilsson, Martin

; APPLICANT: Frykberg, Lars

; APPLICANT: Flock, Jan-Ingmar

; APPLICANT: Lindberg, Martin

; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from

; FILE REFERENCE: Coagulase-Negative Staphylococcus

; CURRENT APPLICATION NUMBER: US/09/147,405B

; PRIOR FILING DATE: 1999-04-11

; PRIOR APPLICATION NUMBER: PCT/SE97/10191

; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: SE 9602496-3

; PRIOR FILING DATE: 1996-06-20

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 15

; LENGTH: 1092

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-147-405B-15

Query Match 34.5%; Score 1628; DB 4; Length 1092;

Best Local Similarity 36.8%; Pred. No. 9.7e-96;

Matches 421; Conservative 141; Mismatches 257; Indels 284; Gaps 32;

QY 5 KKEKHAIRKKSIGVASVLVGTLTIGFGLSSKEADASENSV-----TQSDSASN 52

DB 18 KSNKYAIRKFTVGTASIVIGATILFG-LGHNEAKAEENSVDVKDSDNIDSDSDQSS 76

QY 53 ESKNSDSSSVSAAPKTDDTNV-----SDTKTSNTN---NGETSVANQP 93

DB 77 DEEKNDVINNNQSIINTDDNNQIIKERTNNYVDGIEKRESDRTESTTNVDENEATFLQKTP 136

QY 94 -----AQQETTOSSSTNATEETPTVGEATTTTNCANTPATQS-SNTNAELVN--- 143

DB 137 QDNTHLTETEEVKESSSVESNSSIDTAQPSHTTINREESVQTSNDVSDVDFANSKI 196

QY 144 QTSNETTFNDNTNV-----SSVNSPQNSTNAENVSTQDTSTTEATFSN---NESAP 191

DB 197 KEGNTESKGEENTIEQPNKVKEDSTTSQPSGYNIDKINSQDELLN-LFINEYENKARP 255

QY 192 QSTDASKNOYVNOAVNTSAPRMRFAFSLAAVAADAPAAAGTDITNQLTNVTYGI-DSGTVY 250

```
Db 256 LSTTSAQPSIKRVTVN-----QLAAEQSGNNVHLLIKVTDQ--SITEGYDDSEGVK 304
QY 251 PHQAGVVKLYNFGSPNSAVKGDFTKLTVPKELNNGVTSTAKVPEIMAGD-QVLANGVI 309
Db 305 AHAENLIYDVTFEVDKVKSGDGTMTVDIDKNTVPSDLTSTPIKIKNSGEIATGY 364
QY 310 DS-DGNVIYFTDYVNTKDDVKATILTPAYIDPENVKKTG---NVTLATGIGSTTANKTV 365
Db 365 DNKNQIITVFTDYVDYKNIKAHLKLTYSIDKSVNNNTKLDVEYKTALSS--VNKTI 422
QY 366 LVDYKFKYFNLSKIGTIDQIDKNNVTQRIYVNPNGDNVIAVLTGNLKPNTSNAL 425
Db 423 TVEYQRENRTANLQSMFTNIDTKNHTVEQRIYINPL--RYSAKETNVIISGNGDEGT 480
QY 426 IDQQNTSLIKVYKVDNAADLSESYFV-NPENFEDVNSVNIITFPNPKYKVEVNTPDQIT 484
Db 481 IIDDSTIIKVKYKGNQNLPSNRIDYSEYEDVNDYVAQLGNNNDVNNIFG---NID 536
QY 485 TPYIYVNVGHIDPNSKGD-----LALRSTLYGYSNIIWRMSGMDNEVAFNNGSGSD 537
Db 537 SPYIIKVISKYDPN-KDDVTYTIQQTVMQTTINEYTG--PRTASVDNTIAFSTSSGQ 593
QY 538 G-----IDKXVVP-----EOP-----DEPGEIE 555
Db 594 GDLPPKTYKIGYVWEDVDKGIQNTNDNEXPLSNVLTLTYPDGTSKSVRTDEGKYQ 653
QY 556 -----PIPDSDSDPGSDSGSDS-----NSDSDSG- 582
Db 654 FDLKXGLTYKITFTPEGYFTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTP 713
QY 583 ----- 582
Db 714 KYSLGNVYVYDNTKDGIDDEKIGISGVKVLTKDENGNIISTTTTDENGKYQDNLSGN 773
QY 583 -----SDSTSDSGSDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 631
Db 774 YIVHFKPSGMTQTITDSDGDDDEQADGEEVHTTIDHDDFIDNGYYDDESDSDSDSDS 833
QY 632 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 691
Db 834 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 893
QY 692 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 751
Db 894 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 953
QY 752 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 811
Db 954 DSDSDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1009
QY 812 DSESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 871
Db 1010 DSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1046
QY 872 VPPNPKNGTNASNNKNEAKSKEPLPDGTGSEDE--ANTSLIWCLLASTIGSLLFFRKXEN 929
Db 1047 -----KSTKDKLPDGTANEDYSGKGLTLGLTFLAGLALLGKRRKRN 1088
QY 930 KDK 932
Db 1089 KKK 1091
```

RESULT 8

```
US-09-386-962C-4
; Sequence 4, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: POSTER, Timothy
; TITLE OF INVENTION: POLYPPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; CURRENT FILING DATE: 1999-08-31
```

```
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1742
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-386-962C-4

Query Match 33.5%; Score 1584.5; DB 4; Length 1742;
Best Local Similarity 30.6%; Pred. No. 1e-92;
Matches 452; Conservative 156; Mismatches 285; Indels 583; Gaps 35;

QY 5 KKEKHAIRKKSIGVASVLVGLTILGFLSSKKAADASNSVTSQSDASNE-----SKNSDS 60
Db 28 KVKYSIRKFTVGTASILVGLATMFG-AADNEAKAAEDN--QLESASKEEQKSGRDNENS 84
QY 61 SVS-----AAPKTD-----TNVS-DTTSNT 82
Db 85 KLNQVDDLNGSHSSSEKTNVNNATEVKVBEADTTSDVSKPKANEAVVNESKPKTTEAP 144
QY 83 NNGETSVAQN-----AQOETTQSS-----STNATTEETPTVGEATT-----TTTNOAN 126
Db 145 TVNESIAETPKSTTTQDSTIEKNPDLKDNLSSTTSKESKTDHSTKQACMSINKSN 204
QY 127 -----TPATTQSSNTNABELVNQTSNE-----TTFND----- 153
Db 205 LDTNDSPTQSEKTSQANNSTDNQSAFQKQSDSKPEQKVYKTFNDEFTQDVEHTTK 264
QY 154 -----TNTVSSVNSPQNSTNAENVSTTQDTS-----TEATPSNESAPQSTDAASKOVVN 203
Db 265 LKTPSVSTDSVNDKQDYTRSAVSLGVDNSNETEAITNAVRNLDLKAASRQIIEALIA 324
QY 204 QAV-----NTSAPRMAFSLAAVAADAPAAAGTDITN-- 234
Db 325 EALKKDFSNPDYGVDTPLALNRSQSKNSPHKSASPRMLMSLAA-----EPNSGKNVNDKV 380
QY 235 QLTNVTVGID-----SGTIVPHQAGYVKLYNFGVPSAVKGDFTKLTVPKELNNGVT 289
Db 381 KITNPTLSKNSNNHANNVIMFTSNEQFNKANYELDSDSIKEGDTFTIKYGYIRPGGLE 440
QY 290 STAKVPPIMAGD-QVLANGVIDSDGN-VIYFTFDVVNTKDDVKATLTWPAYIDPN-VKK 346
Db 441 LPAIKTQURSGDGSIVANGVYDKTNTTNTTYFTNTVDQYQNTITGSFDLIATPKRETAID 500
QY 347 TGNVTLATGIGSTTANKTVLVDYKFKYFNLSIKGTIDQIDKNTTYRQTIYVNPSGDN 406
Db 501 NQNYPEVETIANEVVKDFIVD---YGNKKDNTTTAAVANVDNVNKNHNEVVYLLNQNCN 557
QY 407 -----VIAPVLITGNLKN-----TDSNALIDQOQNTSI----- 433
Db 558 PKYAKYFTSVKNGEFIGEVKVEYVTDNMAVDSFNPDLNSNVKDVTSQFAPKVSADGT 617
QY 434 -----KVYKVDNAAD-----LSBSYFWNPENFEDVNSVNIITFPN 468
Db 618 RVDINFARSMAKGYIVTQAVRPTGTCNVVTEYHWTDRDGTNTNDFRGTKSTVTVILN 677
QY 469 PNQ----- 471
Db 678 GSSTAQGDNPVYSLGDYVWLDKKNKGVDDEKGLAGVYVTLKDSNNRELQRTVTTDQSGH 737
QY 472 -----YKVEENTPDDQITTP-----YIVVNGHI----- 495
Db 738 YQFDNLQNGTYVVEFAIPDNTYTPANNSTNDADISDGERDGTGRKVVVAKGTINNADNWT 797
QY 496 -----DPNSKGDIALRSTLYGYNSNIWRMSGMDNE----- 526
Db 798 VDTGYLTFKYNVGVYVWEDNKQIGQDDNEKGISGVKVLTKNKGDTIGTITDTSNGKY 857
QY 527 --VAFNNG-----SGSGDGDID-----KP- 542
```

Db 858 EFTGLENGDYTIEFETPGYPTTKQNSGSDGKSDNGTKITVTYVKDADNKTIDSGFYKPT 917
Qy 543 -----
Db 918 YNLGDYVWEDTNKDGIDQDSEKIGISGVKVTLLKDKNGNAIGTTTTDASGHYQFKGLENGSY 977
Qy 543 -VVEPQP-----
Db 978 TVEFETPGYPTKANSQDITVDSNGITTTGIINGADNLITIDSGFYKTPKYSVGDVYWE 1037
Qy 549 -----DEPGEIEPI-----
Db 1038 DTNKGIDQDNEKIGISGVKVTLLKDKNGNAIGTTTTDENGKYQFONLDSGNVLIHFKEPEG 1097
Qy 560 -----
Db 1098 MTQTANSNGNDEKADGEDVRVITIDHDFSDINGYFDDSDSDSDSDSDSDSDSDSDSDA 1157
Qy 572 DSNDSGSDS 631
Db 1158 DSDSDADS 1217
Qy 632 DS 691
Db 1218 DS 1277
Qy 692 DS 751
Db 1278 DS 1337
Qy 752 DS 811
Db 1338 DS 1397
Qy 812 DS 871
Db 1398 DS 1457
Qy 872 VPENSPKNGTNASNKAKEPLPTGSEDEANT 907
Db 1458 DS 1493

RESULT 9
US-09-386-959-4
; Sequence 4, Application US/09386959
; Patent No. 6703025
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; APPLICANT: FOSTER, Timothy J.
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: MULTICOMPONENT VACCINES
; FILE REFERENCE: P06333USI/BAS
; CURRENT APPLICATION NUMBER: US/09/386.959
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: 60/098,439
; EARLIER FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1742
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-386-959-4

Query Match 33.5%; Score 1584.5; DB 4; Length 1742;
Best Local Similarity 30.6%; Pred. No. 1e-92;
Matches 452; Conservative 156; Mismatches 285; Indels 583; Gaps 35;
Qy 5 KKEKHAIRKXSIGVASVLVGTIGFLLSKKEADASENVTSQSDASNE-----SKSNDSS 60
Db 28 KVKYSIRKFTVGTASILVGTATLMFG-AADNEAKAEDN--QLESASKEQKGRDNE 84

Qy 61 SVS-----AAPKTD-----TNVS-DTKTSST 82
Db 85 KLVQDLNDGSHSEKTTNNNAETVKKVEAPTSDVSKPKANEAIVNNESTKTKTEAP 144
Qy 83 NNGETSAQNP-----AQETTQS-----STNATTEETPTVTEATT-----TTNQAN 126
Db 145 TVNEESIAETPKSTTTQODSTERNPSLKDNLSSSTTSKESKTDHSTKQAOQSTWKS 204
Qy 127 ---TPATTQSNTNAELVNQTSNE-----TFPND----- 153
Db 205 LDTNDSPTQSEKTSQANNDSTNQAPSQLOSKPSEQVYKTKFENDEPTQDVEHTTK 264
Qy 154 ---TNVSSVNSPQNSTNAENVSTTQDTS-----TEATPSNNESAPQSDASNKDVVN 203
Db 265 LKTPSVSDSDVNDKQDYTRSAVASLGVDSNETAITNAVRDNDLJLKAAREQINEAIIA 324
Qy 204 QAV-----NTSAPRMRAFSLAAVAADAPAAAGTDITN-- 234
Db 325 EALKKDFSNPDYGVDTPLALNRQSKNSPHKSASPRVNLMSLAA---EPNSGKNVNDKV 380
Qy 235 QLTNVTVGID-----SGTIVPHQAGVVKLYGSPVNSAVKGTFTKITVPKELNLNGVT 289
Db 381 KITNPTLSLKNHANNVWP*TSNEQNLKANYELDDSIKEGDTFTIKYQYIRPGGLE 440
Qy 290 STAKVPPIMAGD-QVLANGVIDSDGN-VIYTFDYVNTKDDVKATLTMPAVIDPEN-VKK 346
Db 441 LPAIKTLQRLSKDGSIVANGVYDKTNTTTFYFNVDYQYQNTTGSFLLIATPKRETAIKD 500
Qy 347 TGNVTLATGISTTANKTFLVDEYKYGKYNLSIKGTIDQIDKTNNTYRTIYVNPSGDN 406
Db 501 NQNVPEMVTIANEVVKDFIVD---YGNKDNITTAANVDNVNKHNEVYVNLQNNQN 557
Qy 407 ---VIAPVLTGNLKN-----TDSNALIDQNTSI----- 433
Db 558 PKYAKYFTSVKNGEPIGEVKVYEVTDTNAMVDSFNPDLNSSNVKDVTSQFAPKVSADGT 617
Qy 434 ---KVKVVDNAAD-----LSESYFVNPENEDVNSNITFPN 468
Db 618 RVDINFARSANGKKYIVTQAVRPTGNGVYTEWLTRDGTNTNDFYRGTKSTVTYLN 677
Qy 469 PNQ----- 471
Db 678 GSSTAQDNFTYSLGYYVWLDKKNKGVDDEKLAGVYVTLKDSNNRLEQRTVTDQSGH 737
Qy 472 ---YKVEFNTPDQITTP-----YIVVNGHI----- 495
Db 738 YQFDNLQNGTYTVEFAIPDNYTPSPANNSTNDALSDGERDGTGRKVVVAKGTINNADNMT 797
Qy 496 ---DPNSKGDALRSLTYGYSNI.IWRSKSWDNE-- 526
Db 798 VDTGFYLTPKYVGDYVWEDTNKDGIDQDNEKIGISGVKVTLLKKNNGDTIGTTTDSNGY 857
Qy 527 --VAFNNG-----SGSGDGDID-----KP- 542
Db 858 EFTGLENGDYTIEFETPGYPTTKQNSGSDGKSDNGTKITVTYVKDADNKTIDSGFYKPT 917
Qy 543 ----- 542
Db 918 YNLGDYVWEDTNKDGIDQDSEKIGISGVKVTLLKDKNGNAIGTTTTDASGHYQFKGLENGSY 977
Qy 543 -VVEPQP----- 548
Db 978 TVEFETPGYPTKANSQDITVDSNGITTTGIINGADNLITIDSGFYKTPKYSVGDVYWE 1037
Qy 549 -----DEPGEIEPI-----PE- 559
Db 1038 DTNKGIDQDNEKIGISGVKVTLLKDKNGNAIGTTTTDENGKYQFONLDSGNVLIHFKEPEG 1097
Qy 560 -----DSDSDPDSGSDS 571
Db 1098 MTQTANSNGNDEKADGEDVRVITIDHDFSDINGYFDDSDSDSDSDSDSDSDSDSDSDA 1157

APPLICANT: Poster, Timothy J.
APPLICANT: Hook, Magnus A.O.
APPLICANT: Bidhinn, Deirdre Ni
APPLICANT: Perkins, Samuel L.
TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
FILE REFERENCE: P06283US2/BAS
CURRENT APPLICATION NUMBER: US/09/200,650E
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,815
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: 60/098,427
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 3
LENGTH: 930
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-200-650E-3

Query Match 26.3%; Score 1241.5; DB 4; Length 930;
Best Local Similarity 34.9%; Pred. No. 3.7e-71;
Matches 359; Conservative 133; Mismatches 283; Indels 255; Gaps 39;

QY 1 MNMCKK-----EKHAIRKKSIGVASVLVGTLLIGFLLSKKEADASENSVTQSD 48
Db 1 MNMCKKATNPKMIPNPNLNFESIRKYSVGTASILVGTLLIFG-LSGHEAKAAEHTNGELN 59
QY 49 SASNESKNSDSSVSAAPKTDNTNVSPTKTSNTNNGETSAVQNPAAQOETQSSNTNATT 108
Db 60 QSKNETTA--PSENKTKPKVDSRQLK-----NTQTATADQP---KVTMSDS--ATV 104
QY 109 EETPVTGEATTNTNOANTPATTCSSNTNABELVQTSNETTNDNTNVSNSPQNSTN 168
Db 105 KE-----TSSNQSP-----QATANQSTKTSNTNDKSTTYS-----NETD 144
QY 169 AENVSTTQDTSATPSNNESAPOSTDASNKDVVQNAVNTSAPRMRAFSALAAVAADAPAA 228
Db 145 KSNLTQAKDVT-----TPKTTIKPTLNRAVNIVA-----APQQ 181
QY 229 GTDITNLQ--TNVTVGIDSG-----TIVYHQAGVYKLVNFGFVSNVSAVKDGTPTK 278
Db 182 GTNVNDKXVHF-SNIDIALDKGVNQTGKTBFWATSDVLKAKANYTIDDSVKEGDPTFK 241
QY 279 VPKEINLVNVTSTAKVPIM-AGQVILANGVIDSGN-VYFTFDYNTKDDYKATLTMP 336
Db 242 YGYFRFGSVRLPSQTQNLNAQGNIIAKGIYDSTNTTFTYVYDQYTNVRGSEFQV 301
QY 337 AYIDPEN--VKKTG--NVLATIGIGSTTANKVLVDYKGYKFNLSIKGTIDQIDKTN 391
Db 302 AFARKNATDKTAYKMEVTL---GNDTYSEELIIVDYG-----NKAQPLISSTNYIN 351
QY 392 N---TYRQTIYVNPSSGDNVIAPLVTGNL-----KPNTDSNALIDQOQNTSIKVKVDNAAD 443
Db 352 NEDLSRNMTAYVNPQKNTYTKQTFVNTLTGYKFNPA-----KNFKIYEVTDQNG 401
QY 444 LSEGYFVNPENFEDYTSNVTFFPNP-----QYKV-EFNTPDQDITTP 486
Db 402 FVDSFTPTDKLKDVTDFQDIYVSNDRKATFVLMKGQTSNNKQIILQVAYPNSSTD- 460
QY 487 YIVVNGHIDPNSKGLALRSLTYLGYNSNIIWRSMWNEVAFNNGSGSGGDIDK----- 541
Db 461 -----NGKIDYTLTD---KTKY-----SWSNSVNYNGSGSTANGDQKYNLG 500
QY 542 PVWPEQDEPGE-----IEPIEDSDSPGSDSGSDNSD-----SGSDSGSDS--- 585
Db 501 DYVWEDTNKDGQDANEKIGVGVVILKDSNGKELDRITTTDENGKYQFTGLSNGTYSVEF 560
QY 586 -----TSDSGDGAASDSDASDSDASDAS--DSD-----SASD-----SUSDN 625
Db 561 STPAGYTPPTANVGTDDAVSDGLTTTGVIXKADNMTLDSGFKTPKPSLGLDYVWYDSNK 620
QY 626 DSDSDS-----DSDSD-----DSDSD----- 636

Db 621 DGRDSTTEKGIKGVKVTQLQNEKGEVIGTTTETDENGKYRFDNLDSGKYKVFEPAGLITQ 680
QY 637 --SD 694
Db 681 GTNTTDEDDKADAGEVDVVTITDHD-DFTLDNGYVEETSDSDSDSDSDSDSDSDSDSDSDSDSD 739
QY 695 SD 754
Db 740 SD 799
QY 755 SDSE 814
Db 800 SD 859
QY 815 SDSE 871
Db 860 SD 917
QY 872 VPPNSPKNGT 881
Db 918 ---NNSNNGT 924

RESULT 12
US-09-200-650E-5
Sequence 5, Application US/092006650E
Patent No. 6680195
GENERAL INFORMATION:
APPLICANT: Patti, Joseph M.
APPLICANT: Foster, Timothy J.
APPLICANT: Hook, Magnus A.O.
APPLICANT: Bidhinn, Deirdre Ni
APPLICANT: Perkins, Samuel L.
TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
FILE REFERENCE: P06283US2/BAS
CURRENT APPLICATION NUMBER: US/09/200,650E
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,815
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: 60/098,427
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 5
LENGTH: 1315
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-200-650E-5

Query Match 23.5%; Score 1110.5; DB 4; Length 1315;
Best Local Similarity 27.0%; Pred. No. 1.3e-62;
Matches 359; Conservative 175; Mismatches 366; Indels 431; Gaps 42;

QY 1 MNMCKKHAIRKKSIGVASVLVGTLLIGFLG--LSSKEADASENSVTQSDASNESKNSD 58
Db 15 MVSNNLNKESIRKYVTGVTASILVGTLLIFGLGNGQAKAAESTNKELNEATTSASDNQSSD 74
QY 59 SSSVSAAPKTDOT-NVSDTKTSNTNNGETSAVQNPAAQOETQSSNTNATTEPTVTCGA 117
Db 75 KVDNQQLNQEDNTKNDNQKEMVSSQGNETTNGNKLIEKESVQSTTGKV-----EV 126
QY 118 TTTTNTQANTPATTCSSNTNABELVQTSNETTNDNTVSSV---NSPQNSTNAENVST 174
Db 127 STAKSDSQSPKSTNEDLNTKQTSNOEALQDPLQENKSVNVNQPTNEENKVKVDAKTEST 186
QY 175 TQDTSATEPSNNESEA---PQSTDSASNKDVAVQAVNTSAP-----RMRAFSALAAVAADA 225
Db 187 TLNVKSDAIKSNDETLVDNNSNNENNADJI--LPKSTAPKRLNTRMRIAIAVQPSSTEA 244
QY 226 PAAGTDTNOLTNVTVGIDSGTIVYHQAGVYKLVNFGFVSNVSAVKDGTPTKITVPEKELN 285
Db 245 KNVNDLITSTTLTVVDADKNNKIVPAQ-DYLSLSKQITVDDKVKSGDYFTIKYSDTVQV 303

US-09-134-001C-4463

```
Query Match      17.9%; Score 847.5; DB 3; Length 2137;
Best Local Similarity 32.0%; Pred. No. 1.4e-45;
Matches 297; Conservative 149; Mismatches 352; Indels 131; Gaps 21;

QY 33 SKEADASENSVTQ-----SDSASNESKNDSSVSAA-PKDDTINVSDTK 77
Db 1076 SESESDSTSLSESTSLSGSTASTSDASTSDSTSESDSTSESTSLSESTSLSVSDST 1135

QY 78 TSNNTNNGETSVAQ---NPAQQETTQSSSTNATTEETPVTEGATTTTINQANTPATQOS- 133
Db 1136 SASTGESASTSSESESNASTSLGSLSTISDSTSTSTSDASTSSESDSTSTSL 1195

QY 134 SNTNABELVNQSTNETTNDTIVSVSPQNSTN-AENVSTQDTSIEATPSNNEAPQ 192
Db 1196 SESTSTSLSDSTSTSEASTSTSESDSTSESTSLSESTSTSVSDSTSA--STSDSAST 1253

QY 193 STDASNKDVNOAVNTSAPRMEAFSLAAVAADAPAAGT---DITNOLTNVTVGIDSGTTV 249
Db 1254 STSVSDSESASTSISESLSTSVSDSTSTSDASTSTSESDSTSESTSLSESTSTST--- 1310

QY 250 YPHQAGYVKLNGYFVPSNAVKGDTFKITVPKELMLNGVTSTAKVPPINMAGDQVLANGVI 309
Db 1311 -----VSDSTSASTSDASTSTSESESDSASTSLSGSTSTG-----LSDSTG 1352

QY 310 DSDGNVITFDYVNTKDDVKATL-----TMPAYIDPENVKKTGNVTLATGIGSTTANKT 364
Db 1353 TSTSDSASTST---SESDSERASTSLSGSTSTSLSDSTSTSTSDASTSTSVSDSNGAST 1409

QY 365 VLVDYKEYGKFNLSIKGTIDQIDKNTNTYRTIIVNPSGDNVIAPIVLTGNLK--PNTD 421
Db 1410 SL-----SGLSLTSVSDSTSTSDSASTSESDSERASTSLSGSTSTISDST 1459

QY 422 SNALIDQQNTSIKVKVKN-----AADLSESYFVNPFEDVTNSVNIITFNPNOY 472
Db 1460 STSTSDSASTSVSESNSTSTSISESLSTSVSDSTSTSTSDASTSTSVSDSDASTSS 1519

QY 473 KVEFNTPDQIITPTIIVVNGHIDPNSKGDLALRSTLYGYNIIWRSMWDNEVAFNNG 532
Db 1520 SESVSTSDSESTSTSTSDASTST-----TSTSDASTST-----SVSESNSTSTSL 1556

QY 533 SGSGGIDKPVVPEQDFEPIPEDDSDPDGSDGSDNSDSDSDSDSTSDSGSD 592
Db 1557 GSTSTSV-----SDSTSTSDSASTSESDSDSASTSSSESVST 1597

QY 593 SASDSDSASDSDSASDSDSASDSDSDNDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 652
Db 1598 SVSDSTSASTSEASTSTSVSDSNGASTSLSESTSTSLSDSTSMSTSDASTSESDSD 1657

QY 653 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 712
Db 1658 SASTSLSDSTSVSESTSTSTSVASNSTSTSLSDRSLSDSTSTSTSESGSTST 1717

QY 713 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 772
Db 1718 SESDSDSASTSLSESTSTSLSDSTSTSDASTSMVSDSNRASTSLSDSTSTSVSDST 1777

QY 773 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 830
Db 1778 SASTSEASTSTRESESTSASTSLSESTSTSVSDSTSTSTSDASTSTSESDSNGESTSL 1837

QY 831 SDSDSDSASDSDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 888
Db 1838 SESTSTSVSDSTSASTSASTSTSVSDSNGASTSLSGSTSTSV-SDSTSTSTSASTST 1896

QY 889 AKDSKEPLPTGSEDEANTSLIWGLASI 917
Db 1897 TSESE-----DSDSASTSLSGSTSTSI 1917
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:53:31 ; Search time 193.016 Seconds
(without alignments)
2781.239 Million cell updates/sec
Title: US-09-679-643-2
Perfect score: 4725
Sequence: 1 MNMKKEKHAIRKKSIGVAS.....LASIGSLLPRKKENKDKK 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Uniprot.02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4725	100.0	933	2	Q53653	Q53653 staphylococ
2	4553.5	96.4	946	2	Q8XJ1	Q8xj1 staphylococ
3	4542.5	96.1	928	2	Q6GB45	Q6gb45 staphylococ
4	4430	93.8	1029	2	Q6G1K4	Q6gik4 staphylococ
5	4399	93.1	989	2	Q99VJ4	Q99vj4 staphylococ
6	4359	92.3	935	2	Q932C5	Q932c5 staphylococ
7	2495	52.8	881	2	Q93MH7	Q93mh7 staphylococ
8	1810.5	38.3	913	2	Q86476	Q86476 staphylococ
9	1748.5	37.0	907	2	Q8NUL0	Q8nul0 staphylococ
10	1740.5	36.8	905	2	Q6G644	Q6g644 staphylococ
11	1638.5	34.7	877	2	Q99R07	Q99r07 staphylococ
12	1638.5	34.7	877	2	Q7A382	Q7a382 staphylococ
13	1628	34.5	1092	2	Q70022	Q70022 staphylococ
14	1610.5	34.1	1633	2	Q8CMP4	Q8cmp4 staphylococ
15	1584.5	33.5	1733	2	Q9K114	Q9k114 staphylococ
16	1559.5	32.0	873	2	Q6GDH2	Q6gdh2 staphylococ
17	1523.5	32.2	3360	2	Q8XB6	Q8xb6 lactobacill
18	1487	31.5	1056	2	Q8CQ72	Q8cq72 staphylococ
19	1423.5	30.1	1171	2	Q9KWX6	Q9kwx6 staphylococ
20	1414	29.9	1166	2	Q86489	Q86489 staphylococ
21	1395.5	29.5	1385	2	Q99W47	Q99w47 staphylococ
22	1395.5	29.5	1385	2	Q7A780	Q7a780 staphylococ
23	1356.5	28.7	957	2	Q6GBS6	Q6gbs6 staphylococ
24	1356	28.7	1637	1	PLS STAAU	P80544 staphylococ
25	1354.5	28.7	953	2	Q95W48	Q95w48 staphylococ
26	1354.5	28.7	953	2	Q7A781	Q7a781 staphylococ
27	1345	28.5	710	2	Q6R676	Q6r676 staphylococ
28	1345	28.5	710	2	AA500715	AA500715 staphyloc
29	1343.5	28.4	955	2	Q8NXX7	Q8nxx7 staphylococ
30	1325.5	28.1	1365	2	Q6GBS5	Q6gbs5 staphylococ
31	1320.5	27.9	1141	2	Q95W46	Q95w46 staphylococ

32 1320.5 27.9 1141 2 Q932F7 Q932f7 staphylococ
33 1314.5 27.8 1141 2 Q6GBS4 Q6gbs4 staphylococ
34 1314.5 27.8 1141 2 Q8NXX5 Q8nxx5 staphylococ
35 1303 27.6 1893 2 Q8KWM1 Q8kwm1 staphylococ
36 1297.5 27.5 1137 2 Q6GJA6 Q6gja6 staphylococ
37 1286.5 27.2 947 2 Q86487 Q86487 staphylococ
38 1252.5 26.5 1347 2 Q8NXX6 Q8nxx6 staphylococ
39 1155 24.4 906 2 Q6GJA7 Q6gja7 staphylococ
40 1110.5 23.5 1315 2 Q86488 Q86488 staphylococ
41 1055.5 22.3 1698 2 Q8LC00 Q8lc00 staphylococ
42 963 20.4 903 2 Q8Y697 Q8y697 listeria mo
43 957.5 20.3 882 2 Q92AK9 Q92ak9 listeria in
44 883.5 18.7 1253 1 DSPP HUMAN Q9nzw4 homo sapien
45 880 18.6 931 2 Q9K1I3 Q9k1i3 staphylococ

ALIGNMENTS

RESULT 1
Q53653
ID Q53653 PRELIMINARY; PRT; 933 AA.
AC Q53653;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Clumping factor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RX MEDLINE=94224142; PubMed=8170386;
RA McDevitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor (fibrogen receptor
of Staphylococcus aureus.";
RL Mol. Microbiol. 11:237-248(1994)
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (by similarity).
DR EMBL; Z18852; CAI79304.1; -.
DR PIR; S41539; S41539.
DR PDB; 1N67; X-ray; A=202-560.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YSIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF04650; Gram_pos_anchor; 1.
DR TIGRFAMs; TIGR01167; LPXTG signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 933 AA; 97057 MW; EB51A6DE2FF759F4 CRC64;

Query Match 100.0%; Score 4725; DB 2; Length 933;
Best Local Similarity 100.0%; Pred No. 3.3e-194;
Matches 933; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MNMKKEKHAIRKKSIGVASVLVGLTIGFLLSKKEADASENSVTQSDASNESKNDSS 60
Db 1 MNMKKEKHAIRKKSIGVASVLVGLTIGFLLSKKEADASENSVTQSDASNESKNDSS 60
Qy 61 SVSAPKTDITNVSDTKTSSNTNNGETSVAPQAQOETQSSSTNATEETPVTGEATT 120
Db 61 SVSAPKTDITNVSDTKTSSNTNNGETSVAPQAQOETQSSSTNATEETPVTGEATT 120
Qy 121 TTQANTPATTCSSNTNAAELVNQTSNETTNDNTVSSVNSPQNSTNAEVSTTQDTST 180
Db 121 TTQANTPATTCSSNTNAAELVNQTSNETTNDNTVSSVNSPQNSTNAEVSTTQDTST 180
Qy 181 EATPSNNEAPQSDASNKDVVQAVNTSAPRMAFSLAAVAADAPAGTDTTNQLTNVT 240
Db 181 EATPSNNEAPQSDASNKDVVQAVNTSAPRMAFSLAAVAADAPAGTDTTNQLTNVT 240

RT evidence for the rapid evolution of virulence and drug resistance. ;
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).

DR EMBL; BX571856; CAG39851.1; -.
DR InterPro; IPR005877; Gpos YSIRK.
DR InterPro; IPR001899; Gram pos anchor.
DR Pfam; PF00746; Gram pos anchor; 1.
DR TIGRFAMs; TIGR01167; YSIRK_signal; 1.
DR TIGRFAMs; TIGR01167; YSIRK_anchor; 1.
DR TIGRFAMs; TIGR01167; YSIRK_signal; 1.
DR PROSITE; PS50847; GRAM POS ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 1029 AA; 106715 MW; 025E13C6A4C4F020 CRC64;

Query Match 93.8%; Score 4430; DB 2; Length 1029;
Best Local Similarity 86.3%; Pred. No. 1.5e-181;
Matches 888; Conservative 27; Mismatches 18; Indels 96; Gaps 5;

QY 1 MNMKKEKHAIKKKSGVSLVGLTIGFLLSSKEADASENSVTQSDASNEKSDSS 60
DB 1 MNMKKEKHAIKKKSGVSLVGLTIGFLLSSKEADASENSVTQSDASNEKSDSS 60
QY 61 SVSAAKPTDDTNVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGEATT 120
DB 61 SVSAAKPTDDTNVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGEATT 120
QY 121 TTQANTPATTOSSNTAELVNOTSETFNDTNVSSVNSPQNTAENVTSTQDST 180
DB 121 ATQANTPATTOSSNTAELVNOTSETFNDTNVSSVNSPQNTAENVTSTQDST 180
QY 181 EATPSNNEGAPOSTDASNDKVVNOAVNTSAPRMRAFSLAAVAADAPAAAGDTIINLTNVT 240
DB 181 EATPSNNEGAPOSTDASNDKVVNOAVNTSAPRMRAFSLAAVAADAPAAAGDTIINLTNVT 240
QY 241 VGIDSGTTPYHQAGYVKLVGFSVNSAVKGDFTKTIYVPKELNLNGVSTAKVPPIMAG 300
DB 241 VGIDSGTTPYHQAGYVKLVGFSVNSAVKGDFTKTIYVPKELNLNGVSTAKVPPIMAG 300
QY 301 DOVLANGVIDSGNVIYTFDYVNTKDDYKATLTMPAYIDPENVKKTGNVLTATGIGSTT 360
DB 301 DOVLANGVIDSGNVIYTFDYVNTKDDYKATLTMPAYIDPENVKKTGNVLTATGIGSTT 360
QY 361 ANKTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNSGDNVIAPLVLTGNLKPNT 420
DB 361 ANKTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNSGDNVIAPLVLTGNLKPNT 420
QY 421 DSNALIDQONTSIKVKVVDNAADLSSEYFWNPFEDVTNSVNIITPPNPQYKVEFTPD 480
DB 421 DSNALIDQONTSIKVKVVDNAADLSSEYFWNPFEDVTNSVNIITPPNPQYKVEFTPD 480
QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYNLSNIIFWSMWDNEVAFNNGSGGDGID 540
DB 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYNLSNIIFWSMWDNEVAFNNGSGGDGID 540
QY 541 KPVPPEQDPPEGEIPIPEDSDSDP-----GSDSGSDNSDSDSGSDSTSDSGSDSAD 596
DB 541 KPVPPEQDPPEGEIPIPEDSDSDP-----GSDSGSDNSDSDSGSDSTSDSGSDSAD 600
QY 597 SDSASDSASDSASDSDA-----SDSDSASDSDSND 626
DB 601 SDSASDSASDSASDSASDSDSASDSASDSASDSDSTSDSTSDSDSASDSASDSDSD 660
QY 627 SPDS 686
DB 661 SPDS 720
QY 687 SPDSDSDS-----SD 732
DB 721 SPDSDSDSVSD 780
QY 733 SPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSASDSDSDSDSDSDSDSDSDS - 791
DB 791 SPDS

DB 781 SDS 840
QY 792 -----SDSGSD 846
DB 841 SDSDLSD 900
QY 847 SDSSS-----SDSE 864
DB 901 SDSSESDSDSDSDSDSDSDSASDSASDSASDSASDSASDSASDSASDSASDS 960
QY 865 SGSNNNVPPNSPKNGTNASNKNEAKDSKEPLPDTGTSEDEANTSLIWLASIGSLILFR 924
DB 961 SGSNNNVPPNSPKNGTNASNKNEAKDSKEPLPDTGTSEDEANTSLIWLASIGSLILFR 1020
QY 925 RKKNKDKK 933
DB 1021 RKKNKDKK 1029
RESULT 5
Q99VJ4 PRELIMINARY; PRT; 989 AA.
ID Q99VJ4
AC Q99VJ4; DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein A, clumping factor.
GN Name:clfa; Ordered locus names=SA0742;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-U Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus".
RL Lancet 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
DR EMBL; AP003131; BAB41975.1; -.
DR PIR; D89852; D89852.
DR HSSP; Q53653; 1N67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos YSIRK.
DR InterPro; IPR001899; Gram pos anchor.
DR Pfam; PF00746; Gram pos anchor; 1.
DR Pfam; PF04650; YSIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPATG_anchor; 1.
DR PROSITE; PS50847; GRAM POS ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;

Query Match 93.1%; Score 4399; DB 2; Length 989;
Best Local Similarity 88.6%; Pred. No. 3.1e-180;
Matches 876; Conservative 27; Mismatches 30; Indels 56; Gaps 2;

QY 1 MNMKKEKHAIKKKSGVSLVGLTIGFLLSSKEADASENSVTQSDASNEKSDSS 60
DB 1 MNMKKEKHAIKKKSGVSLVGLTIGFLLSSKEADASENSVTQSDASNEKSDSS 60
QY 61 SVSAAKPTDDTNVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGEATT 120
DB 61 SVSAAKPTDDTNVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGEATT 120

```
QY 121 TTNOANTPATQSSNTNAELVNQTSNETTNDNTNTVSSVNSPQNSTNAENVSTQDTST 180
Db 121 TTNOANTPATQSSNTNAELVNQTSNETTNDNTNTVSSVNSPQNSTNAENVSTQDTST 180
QY 181 EATPSNNEAPQSTASNDKDVQVAVNTSAPRMRFAFSLAAVAADAPAAAGTDTITNQLINVT 240
Db 181 EATPSNNEAPQSTASNDKDVQVAVNTSAPRMRFAFSLAAVAADAPAAAGTDTITNQLINVT 240
QY 241 VGDSTGTTVPHQAGYVKNLNGYFVSPNSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAG 300
Db 241 VTDSTGTTVPHQAGYVKNLNGYFVSPNSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAG 300
QY 301 DOVLANGVIDSGNVITFTDYVNTKDDVKATLTPAYIDPENVKKTGNVLTATGISTT 360
Db 301 DOVLANGVIDSGNVITFTDYVNTKDDVKATLTPAYIDPENVKKTGNVLTATGISTT 360
QY 361 ANKTVLDYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAPIVLTGNLKPNT 420
Db 361 ASKTVLDYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAPIVLTGNLKPNT 420
QY 421 DSNALIDQOQNTSIKVYKVDNAADLSESYFVNPENFEDVNTSVNITFPNPQYKVEFTPD 480
Db 421 KSNALIDAKNTDIKVYRVNDNLSESYFVNPENFEDVNTSVNITFPNPQYKVEFTPD 480
QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIIWRSMWMDNEVAFNNGSGSGDID 540
Db 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIIWRSMWMDNEVAFNNGSGSGDID 540
QY 541 KPVPVEQDFPGEIETIPEDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600
Db 541 KPVPVEQDFPGEIETIPEDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600
QY 601 SDSDASDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 622
Db 601 SDSDASDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 660
QY 623 -----SDNPSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 676
Db 623 -----SDNPSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 676
QY 661 SDSDASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 720
Db 661 SDSDASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 720
QY 677 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 736
Db 677 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 736
QY 721 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 780
Db 721 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 780
QY 737 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 796
Db 737 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 796
QY 781 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 840
Db 781 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 840
QY 797 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 853
Db 797 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 853
QY 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Db 841 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
QY 854 -----DSESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 904
Db 854 -----DSESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 904
QY 901 STDTGSDNDSDSNDSSESGSNVPPNPNKNGTNASKNKAKDSKEPLPTGSEDE 960
Db 901 STDTGSDNDSDSNDSSESGSNVPPNPNKNGTNASKNKAKDSKEPLPTGSEDE 960
QY 905 ANTSLIWGLIASIGSLILFRKKENKOKK 933
Db 905 ANTSLIWGLIASIGSLILFRKKENKOKK 933
QY 961 ANTSLIWGLIASIGSLILFRKKENKOKK 989
Db 961 ANTSLIWGLIASIGSLILFRKKENKOKK 989
RESULT 6
Q932C5
ID Q932C5 PRELIMINARY; PRT; 935 AA.
AC Q932C5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibronogen-binding protein.
GN Name=fnb; OrderedLocusNames=SAV0811;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
```

```
OX NCBI_TaxID=158878;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=Mu50 / ATCC 700699;
RA MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Shiba T.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240 (2001).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AP003360; BAB56973.1; -.
DR HSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YSRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSRK_signal; 1.
DR TIGRPFAM; TIGR01167; LpxTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 935 AA; 96949 MW; DC5A2D92CE3BA91C CRC64;

Query Match 92.3%; Score 4359; DB 2; Length 935;
Best Local Similarity 92.5%; Pred. No. 1.5e-178;
Matches 865; Conservative 28; Mismatches 40; Indels 2; Gaps 1;

QY 1 MNMKKEKHAIKKSIGVASVLVGTLLIGFGLLSKEADASENSVTQSDSASNESKSDSS 60
Db 1 MNMKKEKHAIKKSIGVASVLVGTLLIGFGLLSKEADASENSVTQSDSASNESKSDSS 60
QY 61 SVSAAPTDNTNVDTKTSNTNNGETSVACNPAQQTQSSSTNAETETPTVGEATTT 120
Db 61 SVSAAPTDNTNVDTKTSNTNNGETSVACNPAQQTQSSSTNAETETPTVGEATTT 120
QY 121 TTNOANTPATQSSNTNAELVNQTSNETTNDNTNTVSSVNSPQNSTNAENVSTQDTST 180
Db 121 TTNOANTPATQSSNTNAELVNQTSNETTNDNTNTVSSVNSPQNSTNAENVSTQDTST 180
QY 181 EATPSNNEAPQSTASNDKDVQVAVNTSAPRMRFAFSLAAVAADAPAAAGTDTITNQLINVT 240
Db 181 EATPSNNEAPQSTASNDKDVQVAVNTSAPRMRFAFSLAAVAADAPAAAGTDTITNQLINVT 240
QY 241 VGDSTGTTVPHQAGYVKNLNGYFVSPNSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAG 300
Db 241 VTDSTGTTVPHQAGYVKNLNGYFVSPNSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAG 300
QY 301 DOVLANGVIDSGNVITFTDYVNTKDDVKATLTPAYIDPENVKKTGNVLTATGISTT 360
Db 301 DOVLANGVIDSGNVITFTDYVNTKDDVKATLTPAYIDPENVKKTGNVLTATGISTT 360
QY 361 ANKTVLDYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAPIVLTGNLKPNT 420
Db 361 ASKTVLDYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAPIVLTGNLKPNT 420
QY 421 DSNALIDQOQNTSIKVYKVDNAADLSESYFVNPENFEDVNTSVNITFPNPQYKVEFTPD 480
Db 421 KSNALIDAKNTDIKVYRVNDNLSESYFVNPENFEDVNTSVNITFPNPQYKVEFTPD 480
QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIIWRSMWMDNEVAFNNGSGSGDID 540
Db 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIIWRSMWMDNEVAFNNGSGSGDID 540
QY 541 KPVPVEQDFPGEIETIPEDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600
Db 541 KPVPVEQDFPGEIETIPEDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600
```



```
RA Ni Eidhin D., Perkins S., Francois P., Vaudaux P.;  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Newman;  
RA Foster T.J.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
CC an amide bond (by similarity).  
DR EMBL; AJ224764; C:cell surface; IEA.  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR005877; Gpos_Ysirk.  
DR InterPro; IPR001899; Gram_pos_anchor.  
DR Pfam; PF00746; Gram_pos_anchor; 1.  
DR Pfam; PF04650; Ysirk_signal; 1.  
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.  
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.  
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.  
DR Cell wall; Peptidoglycan-anchor; Signal.  
FT SIGNAL 1 44  
FT CHAIN 45 913 Clumping factor B.  
SQ SEQUENCE 913 AA; 97246 MW; 38077C321F8D3E61 CRC64;  
  
Query Match 38.3%; Score 1810.5; DB 2; Length 913;  
Best Local Similarity 43.5%; Pred. No. 1.3e-69; Mismatches 268; Indels 141; Gaps 31;  
Matches 429; Conservative 148;  
  
QY 5 KKEKHAIRKKSIGVASVLVGTIGFGLLSKEADASE--NSVTQSDSASNESKNDSSV 62  
DB 11 KQKYSIRRTVGTTSVIVGATILFG-IGNHQAQASQSDNTTQS-SKNNASADSEKNNM 68  
QY 63 SAAPKTDVTVSDTKTSNTNNGTSTVAQNPAAQOETQSSSTN-ATTEETPTVTEATT 121  
DB 69 IETPQLNTTANDTSDIGANTNSANVSDTKPMSTQTSNTTTTEPASTNETP----OPTAI 124  
QY 122 TNCANTPATQSSNTNAELVNOTSETTENDTNTVSSVNSPQNSTNAENVSTTQDSTE 181  
DB 125 KNOAT--AAKMQDQTVPEQNSQVDNKT--NDANSI-----ATNSE-----LKNSTQL 169  
QY 182 ATPSNESAPOSTDASNKDVQVAVNTSAPMRFAFSLAAVAADAP-----AAGTDITNQ 235  
DB 170 DLP---QSSPQT-----ISNAQGTSKSVRFEAVSLAEPVNAADAKGTNVNDK 218  
QY 236 LTNVTVGIDGTTVPHQAGVVKLVNGSVNSAVKGTFTKIVPKELNLTGVTSTAKVP 295  
DB 219 VTASNFKLEK--TTPDPNQSGNTFMAANFTVTDKVKSGDYF-----TAKLP 262  
QY 296 PIMAGDOVLAVGVID-----SDGNVI-----YTFTDYVNTKD 327  
DB 263 -----DSLGTNGDVVDYNSNNTMPIADIKSTNGDVAVAKATYDILTKYTFVTDVNNKE 317  
QY 328 DVKATLTPAXIDPENVKTKGNVTLATIGSTTANKTVLVDYER-----YGRFYNLSIKGT 383  
DB 318 NINQFSLPFTDRAPAKPSGTVDYANINADIEMFNKITYNSPIAGIDKPNGANSISQ 377  
QY 384 IDQIDKTN--NTYPTQTYVNSP-----GDNVIAPIVLTGNLKNKNTSNALIDQONTSIKVKY 437  
DB 378 IIGVDTASGQNTYKQTVFVNPQKRVGLNTWV--YIKGYQDKIEESSGKVSATDKLRIE 435  
QY 438 VDNAADILSESYFVNP--ENFEDVINSV--NITFPNPQYKVEENTPDDQITTPYIVVNG 493  
DB 436 VNDTSKLSDSYADFNDSNLKEVTDQFKNRIYFHPNVASIKFG-----DITKTVVLVVEG 491  
QY 494 HIDPNSKGDALRSFLYLYN-----SNIWRSMNSMDNEVAFNNGSGSGDIDKPVVPEOPD 549  
DB 492 HYDNTGKN---LKTQVIOENVDPVTRNDRYSIFGNVNNENVRYGGGSADG--DSAVNPKDPT 547  
QY 550 EPGIEPIPE--DSDSDPGSDGSDNSDSDGSDSDTSDSGSDSDSDSDSDSDSDSDSDSDSD 608  
DB 548 GPPVPDPSPDPPEPTPD-----PEPSPDPPEPSPDPSPDPSPDPSPDPSPDPSPDP 589
```

```
609 SDSASDSDSASDSDSDNDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 668  
DB 590 SDGSDSDSDGSDSDSDSESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 649  
QY 669 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 728  
DB 650 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 709  
QY 729 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 788  
DB 710 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 769  
QY 789 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 848  
DB 770 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 829  
QY 849 SSSDSDSDSDSDSDSDSGSNNVVPNSPKNGTNAKNKNEAKDSKEPLPTGTGSEDE-ANT 907  
DB 830 SDSDSDSDSDSR---VTPPNNEOKAFSPKPGEVNHNKSVKQHKTDALPETGDKSENTNA 886  
QY 908 SLIWGLLASIGSLLLFRRKKENKDKX 933  
DB 887 TLFCAMMALIGSLLLFRRKXQDHKEK 912  
  
RESULT 9  
Q8NULO  
ID Q8NULO PRELIMINARY; PRT; 907 AA.  
AC Q8NULO;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Clumping factor B.  
GN Name=clif; OrderedLocusNames=MW2551;  
OS Staphylococcus aureus (strain MW2).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
CX NCBI_TaxID=196620;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=MW2;  
RC MEDLINE=22040717; PubMed=12044378;  
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,  
RA Yamamoto K., Hiramatsu K.;  
RT "Genome and virulence determinants of high virulence community-  
acquired MRSA.";  
RL Lancet 359:1819-1827(2002).  
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
CC an amide bond (by similarity).  
EMBL; AP004831; BAB96416.1; -.  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005618; C:cell wall; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR005877; Gpos_Ysirk.  
DR InterPro; IPR001899; Gram_pos_anchor.  
DR Pfam; PF00746; Gram_pos_anchor; 1.  
DR Pfam; PF04650; Ysirk_signal; 1.  
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.  
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.  
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.  
DR Cell wall; Complete proteome; Peptidoglycan-anchor.  
SQ SEQUENCE 907 AA; 96613 MW; BD951754BCIPE638 CRC64;  
  
Query Match 37.0%; Score 1748.5; DB 2; Length 907;  
Best Local Similarity 42.5%; Pred. No. 5.9e-67;  
Matches 419; Conservative 141; Mismatches 279; Indels 147; Gaps 30;  
  
QY 5 KKEKHAIRKKSIGVASVLVGTIGFGLLSKEADASE--NSVTQSDSASNESKNDSSV 62  
DB 11 KQKYSIRRTVGTTSVIVGATILFG-IGNHQAQASQSDNTTQS-SKNNASADSEKNNM 68  
QY 63 SAAPKTDVTVSDTKTSNTNNGTSTVAQNPAAQOETQSSSTN-ATTEETPTVTEATT 121
```

```
Db 69 IETPOLNTTANDTSDISANTNSANVDSTAKPMSTQTSNTTTEPASTNETP-----QLTAI 124
QY 122 TNOANTPATTTQSSNTNABELVQTSNETTNDNTNTVSSVNSPQNSNAENVSTQDTSTTE 181
Db 125 KQDAT--AAKMQDQVTPQEAQSQVDNKT--NDANSI-----ATNSE-----LKNPQL 169
QY 182 ATPSNNEAPQSTDSASNDKDVNQAVNTSAPRMEAFSLAAVAADAP-----AAGTDITNQ 235
Db 170 DLP---QSSPQT-----ISNAQTSKPSVYTRAVRSLAAVAFVNAADAKGTNVNDK 218
QY 236 LTNVTVGIDSGTTPVPHOAGYKLVNPGFVNSAVKGDTEKTVPKELNLNGVTSTAKVP 295
Db 219 VTAKDFQLEK--TTFDPNQSNTFMAANFTVTGQVKSQDYF-----TAKUP 262
QY 296 PIMAGDQVLANGVID-----SDGNVI-----YTFTDYVNTKD 327
Db 263 -----DSVTGNGDVDSYNSNNMTPIADIIVNDKNEVAKATVDILTKTYTFTVDYVNDKQ 317
QY 328 DVKATLTWPAYIDPENVKTKTGNVLTATGIGSTTANKTVLVVDEK-----YGFYNLSIKGT 383
Db 318 NINGKFSPLFTDRAKAPKSGTYDANINIADEMENNKITNYSSPIAGIDKPGANISSQ 377
QY 384 IDQIDKTN--NTYRQTIYNPS-----GDNVIAPVLVTGNLKPNTDSNALIDQNTSIKYVK 437
Db 378 IIGVDTASQNTYKQTVFVNPQRVLGNTWV--YIKGYQDKIESSGKVSATDTKLRIPE 435
QY 438 VDMAADLSSEYFVNP--ENFEDVTSV--NITFPNPQYKVEFNTPDDQITTPYIVVNG 493
Db 436 VNDTSKLSDSYADPNSNLKEVTQPKDKITYKYQNVASINFG-----DINKTYVVLVEG 491
QY 494 HIDPNSKGDALRSTLYGN-----SNIIWRSMWNEVAFNNGSGSGDGDIDKVPVPEQPD 549
Db 492 HYDKTGKN---LKTQVQENVDPATCKDYISFGWNNENVRYGGGSADG--DSAVNPKDPT 547
QY 550 EFGETIEPIE--DSDDPGSDSGSDNSDSDSGSDSTSDSGSDSASDSDSASDSDSAD 608
Db 548 PGPPVDPPEPPPEPEPPDPPEPTPPPEPPDPPEPPDPPEPPDPPEPPDPPEPPDPPE 607
QY 609 SDSASDSDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 668
Db 608 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 667
QY 669 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 728
Db 668 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 727
QY 729 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 788
Db 728 SESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 787
QY 789 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 848
Db 788 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 834
QY 849 SSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 907
Db 835 -----RVTPPNNEQKAPSNPKGVRHNSKVSQKHKTALPETGDKSENTNA 880
QY 908 SLIWGLLASIGSLLLFREKKENKDKK 933
Db 881 TLFGANMALLSLLLFRRKKQDHKEK 906
RESULT 10
Q6G644
TID Q6G644 PRELIMINARY; PRT; 905 AA.
AC Q6G644;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DE Fibrinogen and keratin-10 binding surface anchored protein.
GN ORNameS=SAS2516;
OS Staphylococcus aureus subsp. aureus MSSA476.
```

```
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]_SEQUENCE FROM N.A.
RC STRAIN=MSSA476;
RA Holden M.T.G., Feil R.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Fellwell I., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Leonard N., Line A., Mayes R., Moulé S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; BX571857; CAG44333.1; -.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs; TIGR01167; YsIRK_signal; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 905 AA; 96411 MW; B8B341A39D800DB8 CRC64;

Query Match 36.8%; Score 1740.5; DB 2; Length 905;
Best Local Similarity 42.5%; Pred. No. 1.3e-66;
Matches 419; Conservative 142; Mismatches 276; Indels 149; Gaps 31;

QY 5 KKEKHAIRKKSIGVASVLVGLIGLLSSKEADASE--NSVTQSDSASNEKSDSSSV 62
Db 11 KONKYSIRRFVTGTTTSVIVGATILFG-ICNHOQAQASEQSNDDTQS--SKNNASADSEKNNM 68
QY 63 SAAPKTDITNYSDDTKTSNTNNGETSVAGNPAQOETQSSSTN-ATTETPTVGTATTT 121
Db 69 IETPOLNTTANDTSDISANTNSANVDSTAKPMSTQTSNTTTEPASTNETP-----QLTAI 124
QY 122 TNOANTPATTTQSSNTNABELVQTSNETTNDNTNTVSSVNSPQNSNAENVSTQDTSTTE 181
Db 125 KQDAT--AAKMQDQVTPQEAQSQVDNKT--NDANSI-----ATNSE-----LKNPQL 169
QY 182 ATPSNNEAPQSTDSASNDKDVNQAVNTSAPRMEAFSLAAVAADAP-----AAGTDITNQ 235
Db 170 DLP---QSSPQT-----ISNAQTSKPSVYTRAVRSLAAVAFVNAADAKGTNVNDK 218
QY 236 LTNVTVGIDSGTTPVPHOAGYKLVNPGFVNSAVKGDTEKTVPKELNLNGVTSTAKVP 295
Db 219 VTAKDFQLEK--TTFDPNQSNTFMAANFTVTGQVKSQDYF-----TAKUP 262
QY 296 PIMAGDQVLANGVID-----SDGNVI-----YTFTDYVNTKD 327
Db 263 -----DSVTGNGDVDSYNSNNMTPIADIIVNDKNEVAKATVDILTKTYTFTVDYVNDKQ 317
QY 328 DVKATLTWPAYIDPENVKTKTGNVLTATGIGSTTANKTVLVVDEK-----YGFYNLSIKGT 383
Db 318 NINGKFSPLFTDRAKAPKSGTYDANINIADEMENNKITNYSSPIAGIDKPGANISSQ 377
QY 384 IDQIDKTN--NTYRQTIYNPS-----GDNVIAPVLVTGNLKPNTDSNALIDQNTSIKYVK 437
Db 378 IIGVDTASQNTYKQTVFVNPQRVLGNTWV--YIKGYQDKIESSGKVSATDTKLRIPE 435
QY 438 VDMAADLSSEYFVNP--ENFEDVTSV--NITFPNPQYKVEFNTPDDQITTPYIVVNG 493
Db 436 VNDTSKLSDSYADPNSNLKEVTQPKDKITYKYQNVASINFG-----DINKTYVVLVEG 491
QY 494 HIDPNSKGDALRSTLYGN-----SNIIWRSMWNEVAFNNGSGSGDGDIDKVPVPEQPD 549
Db 492 HYDKTGKN---LKTQVQENVDPATCKDYISFGWNNENVRYGGGSADG--DSAVNPKDPT 547
```


DR PROSITE; PS0847; GRAM_POS ANCHORING; 1.
KW Cell wall; Peptidoglycan anchor; Signal.
FT SIGNAL 51 Potential.
FT CHAIN 52 1092 Fibrinogen-binding protein.
SQ SEQUENCE 1092 AA; 119292 MW; 6542BC39AAD8B984 CRC64;
Query Match 34.5%; Score 1628; DB 2; Length 1092;
Best Local Similarity 36.8%; Pred. No. 1e-61;
Matches 421; Conservative 141; Mismatches 297; Indels 284; Gaps 32;
QY 5 KKEKHAIRKKSIGVASVLVGTLLIGFLLSKKEADASNSV-----TQSDSASN 52
Db 18 KSNKYAIRKFTVGTASIVGATLLFG-LGHNEAKAEENSVDKVDKNTDDELSDNDQSS 76
QY 53 ESKSNDSVSAAPKTDITNV-----SDTKTSNTN---NGETSVANQP 93
Db 77 DEEKNDVINNNQSIINTDNNQIIKKERTNNYDGIKESDRTESTTNVDNEATFLQKTP 136
QY 94 -----AQQTTSSTNATTEETPVVTEATTTTNTQANTPATQS-SNTNAEELVN--- 143
Db 137 QDNTHLTTEEVEKSSSVESNSSIDTAQPSHTTIKREESVQSDNVDSHVSDFANSKI 196
QY 144 QTSNETFTDNTV-----SSVNSPQNSTNAENVSTTQDTSTTEATPSN---NESAP 191
Db 197 KESNTESGKEENTTEQPNKVKEDSTTQPSGTYTIDEKISNQDELLN-LPINEYENKARP 255
QY 192 QSTDAKNQVNVQAVNTSAPRMFAFSLAAVAADAPAGTDTITNQLTNVTVGI-DSGTIVY 250
Db 256 LSTSAQPSIKRVTVN-----QLAAEQGSNNVHLIKVTDQ--SITEGYDDSEGVIK 304
QY 251 PHQAGYVKLNGFVPSNAVKGDTFKITVPKELNLNGVTSTAKVPPIMAGD-QVLANGVI 309
Db 305 AHDENLIYVTFEVDKVSQDGTMTVDIKNTVPSDLTDSFTIPKIKDMSGEIIAGTY 364
QY 310 DS-DGNVIYTFDYNTKDDVKATLTMPAYIDPENVKKTG---NVTLATGIGSTTANKV 365
Db 365 DNKNKQITVTFDYVDKVENIKAHKLTSYIDKSKVNNNTKLDVEYKLTALSS--VNKTI 422
QY 366 LVDEYKGFKNLSIKGTIDQIDKTNTRYQTIYVNSGDNVIAFVILTNLKPNTDNLAL 425
Db 423 TVEYQRENEARTANLQSMFTNIDPKNHTVEQTIYINEL--RYSAKETNVNISGNGDEGST 480
QY 426 IDQNTSIIKKYKVDNAADLSESFV-FNPEFEDVTNSVNIITFPNPNQYKVEFNTPDQIT 484
Db 481 IIDSTLIKKYKVDGNQLPDSNRIYDYSEYEDVTNDDYAQLGNNDVNIWNGF----NID 536
QY 485 TPYIVVNGHIDPNKGD-----LALRSTLYGYSNIIWRSMWNEVAFNNGSGSGD 537
Db 537 SPYIIKVISKYDEN-KDDYTTIQQVTMTQTTINEYTGEB--FRTASYDNTIAFSTSGQGQ 593
QY 538 G-----IDKPVVP-----EQP-----DEPGEIE 555
Db 594 GDLPEPTYKIGYVWEDVDKGIQNTNNDNEKPLSNVLVLTLPDGTSKSVRTDEDGKYQ 653
QY 556 -----PIPEDSDSDPGSGSGSDS-----NSDSDSGD----- 582
Db 654 FDGLKNGLYKITPETEGYPTLKHSGTNPALDSEGNVWVWITNGQDDMTIDSGFYQTP 713
QY 583 ----- 582
Db 714 KYSLGNVWYDYNKDGIIQGDDEKIGSGVKVTKLDENGNIISTTTTDENGKYQFDNLNSGN 773
QY 583 -----SDSTSDSGSDASDSASDSASDSASDSASDSASDSASDSASDSASDS 631
Db 774 YIVHFDKPSGMGTQTTFSDGDDDEQADGEEVHVHTIDHDQDFSIDNGYVYDSDSDSDS 833
QY 632 DS 691
Db 834 DS 893
QY 692 DS 751
Db 894 DS 953

QY 752 DS 811
Db 954 DSDSDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1009
QY 812 DS 871
Db 1010 DSVSDSDSDSDSDSGSDS 1046
QY 872 VPNSPKNGTNAKNKAEKADSKPELPTGSEDE--ANTSILWGLLASIGSILLIFRRYKEN 929
Db 1047 -----KSTKDKLPDTGANEDYGSKGTLLGLTFLFAGLGLALLGKRKRNR 1088
QY 930 KDK 932
Db 1089 KNK 1091
RESULT 14
Q8CMP4
ID Q8CMP4 PRELIMINARY; PRT; 1633 AA.
AC Q8CMP4; 2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
GN OrderedLocusNames=SE2395;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
Qian Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RA "Genome-based analysis of virulence genes in a non-biofilm-forming
ST Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593 (2003).
DR EMBL; AE016752; AA006038.1;
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR008877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B_4.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 1633 AA; 174769 MW; DAA8537B8CD288BD CRC64;
Query Match 34.1%; Score 1610.5; DB 2; Length 1633;
Best Local Similarity 29.5%; Pred. No. 9.4e-61;
Matches 480; Conservative 153; Mismatches 282; Indels 713; Gaps 39;
QY 5 KKEKHAIRKKSIGVASVLVGTLLIGFLLSKKEADASNSVTSQSDASNE-----SKSDSDS 60
Db 19 KVKYKIRKFTVGTASIVGATLLFG-AADNEAKAEDN-QLESASKEEQKSGSRDNESS 75
QY 61 SVS-----AAPKTD-----TNVS-DTKTSNT 82
Db 76 KLNQVDLNGSHSEKTTNNVNAETVKKVEAPTTSDVSKPKANEAVVNTNESTPKTKTEAP 135
QY 83 NNGETSVANQP-----AQQTTSQSS-----STNATTEETPTVTEATT---TTTNOAN 126
Db 136 TVNEESAETPKYTTTQDDSTKNNPNSLKNLNSSTTSKESKDEHSTTQQAQMSKNS 195
QY 127 ---TPATTSQNTNNAEELVNQTSNE-----TTFND----- 153

Db 136 TVNEESIAETPKTSTTQODSTKNNPSLKDNLNSSSTTSKESKTDBHSTKQAOQWSTNKS 195
Qy 127 ---TPATTQSSNTNAEELVNOTSNE-----TTND----- 153
Db 196 LDTNDPTQSEKTSQANNDSTNQAPSQOLDSKPSEQKVYTKFNDEPTQVEHTTK 255
Qy 154 ---TNTVSSVNSPQNSTNAENVSTTQDTS---TEATPSNNESAPOSTDASNKDVVN 203
Db 256 LKTPSVSTSSVNDKQDTRSAVASLGVDSNEEAITNAVRDNLDLKAAAREQINEALIA 315
Qy 204 QAV-----NTSAPRMRAFSLAAVAADAPAAAGTDITN-- 234
Db 316 EALKKDFNSPDYGVDTPLALNRQSOKNSPKHSASPRNMLSLAA---EPNSGKNVNDKV 371
Qy 235 OLNTNVLGID-----SGTIVYHQAGVVKLYNGFSPNSAVKGTFFKLTVPKELNLNGVT 289
Db 372 KITNPTLSLNKSNHANNVWTPTSNEOFNLKANYELDDSIKEGFTFKYQGVIRPGGLE 431
Qy 290 STAKVPPIMAGD-QVLANGVIDSGN-VIYTFDYVNTKDDVKATLTPAYIDPEN-VKK 346
Db 432 LPAIKTQLRSKDSIVANGVYDKTNTTTFYFNVQYQNIITGSPDLIATPKRETAIKD 491
Qy 347 TGNVTLATGISTANKTVLVDYKFKYNLSIKGTIDQIDKTNNTYROTIVYNPBGDN 406
Db 492 NQNPMEVTTIANEVKKDFIVD---YGNKKDNTTTAAVANVDVNNKHNEVVYLNQNNQN 548
Qy 407 ---VIAPVLTGNLKEN-----TDSNALIDQOQNTS1----- 433
Db 549 PKYAKFSTVKNGEFIPGEVKVVEVDTNAWVDSFNDPLNSSNVKDVTSQFAPKVSADGT 608
Qy 434 ---KVKYVDNAAD-----LSESYFVNPFENFEDVTSNYNITFPN 468
Db 609 RVDINPARSMANGKYYIVTQAVRPTGNGVYTEWLTFRDGTNTINDFYRGTKSTTVTYLN 668
Qy 469 PNQ----- 471
Db 669 GSSTAQDNPTYSLGDIYVWLDKKNKGVDQDDDEKGLAGVYVTLKDSNNRELQRTVTTQSGH 728
Qy 472 ---YKVEFNTPDQITTP-----YIVVNVNGHI----- 495
Db 729 YQFDNLQNGTYTVBEFALPDNTPSPANNSTNDAIDSGERDGTGRKVVVAKGTINNADNMT 789
Qy 496 ---DPNSKGLALRSTLYGYNSNIIWRMSWDNE-- 526
Db 789 VDTGFIYLPKXNVGDIYVWEDTNKDGIQDDNEKGLSGVKVTLKKNKGDTIGTTITDSNGKY 848
Qy 527 --VAPNNG-----SGSGGID-----KP- 542
Db 849 EFTGLENGDYTIEFETPEGTPTKQNSGSDGKDSNGTKTTVTVKADANKTIDSGFYKPT 908
Qy 543 ----- 542
Db 909 YNLGDIYWEDTNKDGIQDDSEKGISGVKVTLLKKNAGNAIGTTTTDASGHYQFKGLENGSY 968
Qy 543 -VVPEQP----- 548
Db 969 TVBEFETSGYTPTKANSQDITVDSNGITTTGIINGADNLIDSGFYKPKYSVGDIYWE 1028
Qy 549 -----DEPGEIEPI-----PE- 559
Db 1029 DTNKGDIQDQNEKGISGVKVTLLKDEKGNIISTTTTDDENGKYQFDNLDSGNIIHFKEPGE 1088
Qy 560 -----DSDSDPGSDSGS 571
Db 1089 MTQTTANSNDDEKADGEDVRVITIDHDFSIDNGYFDDSDSDSDSDSDSDSDA 1148
Qy 572 DSNDSGSDSGSDSTSGSDSGSDSDSASDSDSASDSDSASDSDSDSDSDSDSDSDSDSDS 631
Db 1149 DSDSDADS 1208
Qy 632 DS 691
Db 1209 DS 1268

Qy 692 DS 751
Db 1269 DSDSDADS 1328
Qy 752 DS 811
Db 1329 DS 1388
Qy 812 DSESDS 871
Db 1389 DSDA 1448
Qy 872 VPNSPKNGTNAKNKNEAKSKEPLPTGTSEDEANT 907
Db 1449 DS 1484

Search completed: November 3, 2004, 19:10:15
Job time : 204.016 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:48:10 ; Search time 100.005 Seconds
(without alignments)
1894.005 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550
Perfect score: 2702
Sequence: 1 VGLTIGFLLSSREADSEN.....NGSGSGDIDKPVVPEQDE 528

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2702	100.0	933	3	AA558435	AA558435 Staphyloc
2	2702	100.0	933	4	AB69508	AB69508 Staphyloc
3	2694	99.7	933	6	ABJ18947	ABJ18947 Pathogen
4	2690	99.6	927	6	ABM72221	ABM72221 Staphyloc
5	2690	99.6	936	2	AAW89801	AAW89801 Staphyloc
6	2621	97.0	520	5	AAE22262	AAE22262 Staphyloc
7	2466	91.3	935	6	ABU16402	ABU16402 Protein e
8	2450	90.7	1021	4	AAU33975	AAU33975 Staphyloc
9	2450	90.7	1021	4	AAU38951	AAU38951 Staphyloc
10	2408	89.1	496	5	AAU75490	AAU75490 S. aureus
11	2408	89.1	496	6	ADA89664	ADA89664 Staphyloc
12	2408	89.1	496	6	ADA89663	ADA89663 Staphyloc
13	1727	63.9	331	5	AAE22263	AAE22263 Staphyloc
14	1727	63.9	345	2	AAW31555	AAW31555 Fibronect
15	520	19.2	961	6	ABU15854	ABU15854 Protein e
16	515	19.1	767	4	AAU34403	AAU34403 Staphyloc
17	515	19.1	767	4	AAU37547	AAU37547 Staphyloc
18	515	19.1	940	6	ABJ19111	ABJ19111 Pathogen
19	515	19.1	948	6	ADA89470	ADA89470 Staphyloc
20	515	19.1	948	6	ABJ18923	ABJ18923 Pathogen
21	515	19.1	948	6	ABM72536	ABM72536 Staphyloc
22	512	18.9	940	2	AAU07070	AAU07070 Fibronect
23	494.5	18.3	1166	2	AAU08643	AAU08643 S. aureus
24	494.5	18.3	1166	6	ABJ18982	ABJ18982 Pathogen
25	494	18.3	1141	6	ABU42327	ABU42327 Protein e

26	472	17.5	1092	2	AAW41602	AAW41602 Staphyloc
27	472	17.5	1092	7	ABM79019	ABM79019 Staphyloc
28	471.5	17.5	930	5	ABP40469	ABP40469 Staphyloc
29	471.5	17.5	991	3	AA583171	AA583171 Cell wall
30	471.5	17.5	991	3	AAU70120	AAU70120 Staph. ep
31	468	17.3	1001	4	AAU37093	AAU37093 Staphyloc
32	467	17.3	567	6	ADA89555	ADA89555 Staphyloc
33	467	17.3	877	6	ADA89539	ADA89539 Staphyloc
34	467	17.3	877	6	ABU42504	ABU42504 Protein e
35	467	17.3	877	6	ABM72702	ABM72702 Staphyloc
36	467	17.3	913	6	ABJ18917	ABJ18917 Pathogen
37	465.5	17.2	892	6	ABU42557	ABU42557 Protein e
38	465.5	17.2	892	8	ADO84849	ADO84849 S. epiderm
39	464	17.2	918	2	AAU08640	AAU08640 S. aureus
40	460.5	17.0	1018	4	AAU37245	AAU37245 Staphyloc
41	460.5	17.0	1018	4	AAU34301	AAU34301 Staphyloc
42	460.5	17.0	1018	6	ABJ18922	ABJ18922 Pathogen
43	456	16.9	978	4	AAU33960	AAU33960 Staphyloc
44	455.5	16.9	549	7	ABM79020	ABM79020 Staphyloc
45	455.5	16.9	560	7	ABM79015	ABM79015 Staphyloc

ALIGNMENTS

RESULT 1
AA558435
ID AA558435 standard; protein; 933 AA.
XX AC AA558435;
XX 27-MAR-2000 (first entry)
XX DE Staphylococcus aureus fibrinogen binding ClfA protein.
XX ClfA; fibrinogen binding protein; bacterial colonisation;
KW indwelling medical device; staphylococcal infection.
XX OS Staphylococcus aureus.

XX Key Location/Qualifiers
FT Peptide 1..39
FT Region 40..559 /note= "Signal peptide"
FT Region 332..550 /note= "Region A"
FT Region /note= "Fibrinogen-binding region"
FT Region 560..867 /note= "Region R"
FT Region 896..900 /note= "Gram positive wall-associated consensus motif"

US6008341-A.

28-DEC-1999.

22-AUG-1994; 94US-00293728.

22-AUG-1994; 94US-00293728.

(QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

Foster TJ, McDevitt DL;

WPI; 2000-096389/08.

N-PSDB; AAZ55832.

Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus

aureus, useful for treatment or prevention of infections.

Disclosure; Fig 2A-1-4; 35pp; English.

This sequence represents the Staphylococcus aureus fibrinogen-binding

app. 1.00

CC protein, ClfA. ClfA is an important receptor involved in *S. aureus*
CC colonisation of indwelling medical devices (e.g., catheters, artificial
CC heart valves). Shortly after implantation, the surfaces of medical
CC devices become coated with host plasma and matrix proteins such as
CC fibrinogen and fibronectin, and there is considerable evidence to suggest
CC that bacterial adherence to fibrinogen/fibrin is important in the
CC initiation of device-related infection. The fibrinogen-binding region of
CC ClfA is thought to reside between residues 332 and 550 in a region R)
CC designated A. The protein also contains a repeated region (region R)
CC comprising 154 repeats of the dipeptide Ser-Asp, and the C-terminus
CC contains features present in surface proteins of other Gram positive
CC bacteria that are responsible for anchoring the protein to the cell wall
CC and cell membrane. ClfA, or its fragments, may be used to block *S. aureus*
CC colonisation of wounds, to prevent adherence of *S. aureus* to indwelling
CC medical devices, as vaccines to protect against *S. aureus* infection
CC (e.g., mastitis in ruminants), to raise specific antibodies, and for
CC diagnosis (by agglutination or immunoassay). The specific antibodies are
CC used for passive immunisation, to block infection of wounds or adhesion
CC of *S. aureus* and for diagnosis. Nucleotides encoding ClfA and its
CC fragments may be used as diagnostic probes
XX

XX Sequence 933 AA;

Query Match 100.0%; Score 2702; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 3.7e-140;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VCTLIGFGLLSKEADASENSVTQSDASNESKSDSSVSAPKTDNTNVSDTKTSNT 60
DB 23 VCTLIGFGLLSKEADASENSVTQSDASNESKSDSSVSAPKTDNTNVSDTKTSNT 82
QY 61 NNGETSVAQNPAQOETQSSSTNATTEETPVGTGEATTTTNOANTPATQSSNTNAEELV 120
DB 83 NNGETSVAQNPAQOETQSSSTNATTEETPVGTGEATTTTNOANTPATQSSNTNAEELV 142
QY 121 NOTSETTFNDTNTVSSVNSPQNSTNAENVTQTSTTEATPSNNEAPQSTDAKNQV 180
DB 143 NOTSETTFNDTNTVSSVNSPQNSTNAENVTQTSTTEATPSNNEAPQSTDAKNQV 202
QY 181 NOAVNTSAPRMRAFSLAAVAADAPAGTDITNQLTNVTVGIDSGTTPVPHQAGYKVLNG 240
DB 203 NOAVNTSAPRMRAFSLAAVAADAPAGTDITNQLTNVTVGIDSGTTPVPHQAGYKVLNG 262
QY 241 FSVNPSAVKGDFTFKITVPKELNLTGTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 300
DB 263 FSVNPSAVKGDFTFKITVPKELNLTGTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 322
QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKVLDYKYGKFPYLSIKG 360
DB 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKVLDYKYGKFPYLSIKG 382
QY 361 TIDQIDKTNNTYRQTIYVNPFGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAA 420
DB 383 TIDQIDKTNNTYRQTIYVNPFGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAA 442
QY 421 DLSESYFVNPENFEDVNTSVNITFPNPNQYKVEFNTPDQITTPYIVVNGHIDPNSKGD 480
DB 443 DLSESYFVNPENFEDVNTSVNITFPNPNQYKVEFNTPDQITTPYIVVNGHIDPNSKGD 502
QY 481 LAIRSTLYGNSNIIWRSMNSWNEVAFNNGSGSGDIDKPVVPEQDPE 528
DB 503 LAIRSTLYGNSNIIWRSMNSWNEVAFNNGSGSGDIDKPVVPEQDPE 550

RESULT 2

AAB69508
ID AAB69508 standard; protein; 933 AA.

XX
XX AAB69508;

XX
XX 23-APR-2001 (first entry)

XX
XX Staphylococcus aureus ClfA protein.

XX
KW Staphylococcus aureus; clfA; antibiotic; vaccine;
KW fibrinogen binding protein; bacterial infection; mastitis.
XX
OS Staphylococcus aureus.
PN US6177084-B1.
XX
PD 23-JAN-2001.
XX
PF 19-OCT-1999; 99US-00421869.
XX
PR 22-AUG-1994; 94US-00293728.
XX
PA (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
PI Foster TJ, Medevitt DL;
XX
DR WPI; 2001-181608/18.
DR N-PSDB; AAF58593.
XX
PT Novel Staphylococcus aureus fibrinogen binding protein useful as vaccine
PT to prevent infection, promote wound healing, block adherence to
PT indwelling medical devices and for diagnosing staphylococcus aureus
PT infection.
XX
PS Claim 5; Fig 2; 30pp; English.
XX
CC The present sequence is a novel Staphylococcus aureus fibrinogen binding
CC protein. It is useful as a vaccine to protect against human and animal
CC infections caused by *S. aureus*, such as against mastitis, to block
CC *S. aureus* from colonising and infecting a wound, to block adherence of
CC *S. aureus* to indwelling medical devices such as catheters, replacement
CC heart valves and cardiac assist devices. The protein can be used to
CC diagnose bacterial infections. Polyclonal and monoclonal antibodies
CC raised against the fibrinogen binding protein are useful for passively
CC immunising against infections caused by *S. aureus*, to prevent infection of
CC a wound and to diagnose bacterial infections
XX
SQ Sequence 933 AA;

Query Match 100.0%; Score 2702; DB 4; Length 933;
Best Local Similarity 100.0%; Pred. No. 3.7e-140;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VCTLIGFGLLSKEADASENSVTQSDASNESKSDSSVSAPKTDNTNVSDTKTSNT 60
DB 23 VCTLIGFGLLSKEADASENSVTQSDASNESKSDSSVSAPKTDNTNVSDTKTSNT 82
QY 61 NNGETSVAQNPAQOETQSSSTNATTEETPVGTGEATTTTNOANTPATQSSNTNAEELV 120
DB 83 NNGETSVAQNPAQOETQSSSTNATTEETPVGTGEATTTTNOANTPATQSSNTNAEELV 142
QY 121 NOTSETTFNDTNTVSSVNSPQNSTNAENVTQTSTTEATPSNNEAPQSTDAKNQV 180
DB 143 NOTSETTFNDTNTVSSVNSPQNSTNAENVTQTSTTEATPSNNEAPQSTDAKNQV 202
QY 181 NOAVNTSAPRMRAFSLAAVAADAPAGTDITNQLTNVTVGIDSGTTPVPHQAGYKVLNG 240
DB 203 NOAVNTSAPRMRAFSLAAVAADAPAGTDITNQLTNVTVGIDSGTTPVPHQAGYKVLNG 262
QY 241 FSVNPSAVKGDFTFKITVPKELNLTGTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 300
DB 263 FSVNPSAVKGDFTFKITVPKELNLTGTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 322
QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKVLDYKYGKFPYLSIKG 360
DB 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKVLDYKYGKFPYLSIKG 382
QY 361 TIDQIDKTNNTYRQTIYVNPFGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAA 420
DB 383 TIDQIDKTNNTYRQTIYVNPFGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAA 442

Appendix

QY 421 DLSESYFVNPFENFEDVTNSNIIFPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGD 480
 DB 443 DLSESYFVNPFENFEDVTNSNIIFPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGD 502
 QY 481 LALRSTLYGYNIIWRSMWDNEVAFNNGSGDGDIDKFPVPEQDPE 528
 DB 503 LALRSTLYGYNIIWRSMWDNEVAFNNGSGDGDIDKFPVPEQDPE 550

RESULT 3
 ABJ18947
 ID ABJ18947 standard; protein; 933 AA.
 XX
 AC ABJ18947;
 XX
 DT 06-MAR-2003 (first entry)
 XX
 DE Pathogen specific antigen related staphylococcal protein SEQ ID No 93.
 XX
 KW Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
 KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
 KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
 KW autoimmune disease; HIV; hepatitis.
 XX
 OS Staphylococcus sp.
 XX
 PN WO200259148-A2.
 XX
 PD 01-AUG-2002.
 XX
 XX 21-JAN-2002; 2002WO-EP000546.
 XX
 XX 26-JAN-2001; 2001AT-00000130.
 PR
 XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
 PA
 XX Meinke A, Nagy E, Von Ahseu M, Klade C, Henics T, Zauner W;
 PI Minik DB, Vytvyska O, Etz W, Dryla A, Weichhart T, Hafner M;
 PI Tempelmeier B;
 XX
 XX WPI; 2003-075410/07.
 XX
 XX Identifying, isolating and producing hyperimmune serum-reactive antigens
 XX from a pathogen, for preparing vaccine or medicament for treating or
 XX preventing e.g. staphylococcal infections, comprises providing antibody
 XX preparation.
 XX
 XX Example 7; Page 160; 252pp; English.
 XX
 CC The invention relates to a novel method for identifying, isolating and
 CC producing hyperimmune serum-reactive antigens from a pathogen, tumour,
 CC allergen, a tissue or host prone to auto-immunity, where the antigens are
 CC used in a vaccine, comprises providing antibody preparation from a plasma
 CC pool of a type of animal, or individual sera with antibodies against the
 CC specific pathogen, tumour, allergen, tissue or host prone to auto-
 CC immunity. The hyperimmune serum-reactive antigens comprising any of the
 CC 62 sequences of 53-2261 amino acids fully defined in the specification,
 CC or their hyperimmune fragments are useful for the manufacture of a
 CC pharmaceutical preparation, particularly a vaccine against staphylococcal
 CC infections or colonisation against S. aureus or S. epidermidis. The
 CC preparation of antibodies is useful for the manufacture of a medicament
 CC for treating or preventing staphylococcal infections or colonisation
 CC against S. aureus or S. epidermidis. The antibody preparations may also
 CC be used for diagnostic and imaging purposes. Other conditions that can be
 CC treated include cancer, autoimmune diseases or infections caused by viral
 CC (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
 CC sequence represents a staphylococcal protein relating to the method for
 CC identifying and producing pathogen specific antigens of the invention
 XX
 SQ Sequence 933 AA;
 Query Match 99.7%; Score 2694; DB 6; Length 933;
 Best Local Similarity 99.8%; Pred. No. 1e-139;

Matches 527; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VGLIFGGLLSKEADASENSVTQSDASNESKNSDSSVSAAAPKTDITNVSDTKTSNT 60
 DB 23 VGLIFGGLLSKEADASENSVTQSDASNESKNSDSSVSAAAPKTDITNVSDTKTSNT 82
 QY 61 NNGETSAQNPAQOETTQSSSTNAITTEETPVTGEATTTTINQANTPATTOSSTNAEELV 120
 DB 83 NNGETSAQNPAQOETTQSSSTNAITTEETPVTGEATTTTINQANTPATTOSSTNAEELV 142
 QY 121 NQTSNETTNDTNTVSSVNSPQNSTNAENVSTTQDTSSTATPSNNSAPOSTDASNKDVV 180
 DB 143 NQTSNETTNDTNTVSSVNSPQNSTNAENVSTTQDTSSTATPSNNSAPOSTDASNKDVV 202
 QY 181 NOAVNTSAPRMRAFSLAAVAADAPAAAGTDITNLTNTVVGIDSGTTTYPHQAQGVVKNLYG 240
 DB 203 NOAVNTSAPRMRAFSLAAVAADAPAAAGTDITNLTNTVVGIDSGTTTYPHQAQGVVKNLYG 262
 QY 241 FSPNSAVKGDITFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSGNVIYTFDY 300
 DB 263 FSPNSAVKGDITFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSGNVIYTFDY 322
 QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKYGKFNLSIKG 360
 DB 323 VNTKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKYGKFNLSIKG 382
 QY 361 TIDQIDKNTNYRTIYVNPFGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIVKVKVDNAA 420
 DB 383 TIDQIDKNTNYRTIYVNPFGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIVKVKVDNAA 442
 QY 421 DLSESYFVNPFENFEDVTNSNIIFPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGD 480
 DB 443 DLSESYFVNPFENFEDVTNSNIIFPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGD 502
 QY 481 LALRSTLYGYNIIWRSMWDNEVAFNNGSGDGDIDKFPVPEQDPE 528
 DB 503 LALRSTLYGYNIIWRSMWDNEVAFNNGSGDGDIDKFPVPEQDPE 550

RESULT 4
 ABM72221
 ID ABM72221 standard; protein; 927 AA.
 XX
 AC ABM72221;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Staphylococcus aureus protein #1461.
 XX
 KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 KW enzymatic assay; antibiotic target.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200294868-A2.
 XX
 PD 28-NOV-2002.
 XX
 XX 27-MAR-2002; 2002WO-IB002637.
 PF
 XX 27-MAR-2001; 2001GB-00007661.
 PR
 XX (CHIR-) CHIRON SPA.
 XX
 XX Masignani V, Mora M, Scarselli M;
 PI WPI; 2003-120285/11.
 XX
 DR N-PSDB; ACF73781.
 XX
 PT New Staphylococcus aureus protein, useful as a vaccine for treating or
 PT preventing Staphylococcal infection, specifically an infection caused by
 PT S. aureus, e.g. sepsis.
 XX

PS Claim 1; SEQ ID NO 2922; 49pp; English.

XX The invention relates to novel genes and encoded proteins from

CC Staphylococcus aureus. A composition comprising the S. aureus protein, a

CC nucleic acid encoding the protein, or an antibody to the protein, is

CC useful as a pharmaceutical, particularly as a vaccine for treating or

CC preventing infection due to Staphylococcus bacteria, specifically an

CC infection caused by S. aureus. The composition is particularly useful for

CC treating or preventing sepsis in a patient. The composition can also be

CC used for diagnostics. The protein is also used in an assay for enzymatic

CC studies and as a target for antibiotics. This sequence represents one of

CC the novel S. aureus proteins of the invention

XX Sequence 927 AA;

Query Match 99.6%; Score 2690; DB 6; Length 927;

Best Local Similarity 99.6%; Pred. No. 1.7e-139;

Matches 526; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGTLLGFLSSKEADASENSVTQSDASNESKNDSSVSAAPKTDITNVSDTKTSNT 60

Db 23 VGTLLGFLSSKEADASENSVTQSDASNESKNDSSVSAAPKTDITNVSDTKTSNT 82

QY 61 NNGETSAQNPAQOQTQSSSTNATTEPTVTGEATTTTNOANTPATQSSNTNAEELV 120

Db 83 NNGETSAQNPAQOQTQSSSTNATTEPTVTGEATTTTNOANTPATQSSNTNAEELV 142

QY 121 NOTSETTNDNTVSSVNSPQNSINAENVSTQDTSATPSNNESAPQSDASNKQVY 180

Db 143 NOTSETTNDNTVSSVNSPQNSINAENVSTQDTSATPSNNESAPQSDASNKQVY 202

QY 181 NQAVNTSAPRMAFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYKLVNG 240

Db 203 NQAVNTSAPRMAFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYKLVNG 262

QY 241 FSVNPSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 300

Db 263 FSVNPSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 322

QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKTVLDVDEYKGYFYNLSIKG 360

Db 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKTVLDVDEYKGYFYNLSIKG 382

QY 361 TIDQIDKTNNTYRQTIYVNPNGDNVIAVLGTGNLKPNTDSNALIDQNTSIKVKVDNAA 420

Db 383 TIDQIDKTNNTYRQTIYVNPNGDNVIAVLGTGNLKPNTDSNALIDQNTSIKVKVDNAA 442

QY 421 DLSESYFVNPENFEDVNTSVNITFPNPNQYKVEFNTPDQITTPYIVVNGHIDPNSKGD 480

Db 443 DLSESYFVNPENFEDVNTSVNITFPNPNQYKVEFNTPDQITTPYIVVNGHIDPNSKGD 502

QY 481 LALRSTLYGNSNIIWRSMWNEVAFNNGSGSGDIDKPVVPEQFDE 528

Db 503 LALRSTLYGNSNIIWRSMWNEVAFNNGSGSGDIDKPVVPEQFDE 550

RESULT 5

AAW89801

ID AAW89801 standard; protein; 936 AA.

XX AC AAW89801;

DT 16-MAR-1999 (first entry)

XX Staphylococcus aureus protein SEQ ID #5249.

XX Computer readable medium; vaccine; S. aureus infection; immunodetection;

XX cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;

XX skin infection; surgical wound infection; scalded skin syndrome;

XX toxic shock syndrome.

XX Staphylococcus aureus.

XX

PN EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.

XX 05-JAN-1996; 96US-0009861P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Claim 23; Page 3255-3258; 3271pp; English.

XX This sequence represents a Staphylococcus aureus protein sequence of the

CC invention. The DNA sequences encoding the S. aureus proteins are recorded

CC on a computer readable medium, preferably selected from a floppy or hard

CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.

CC Homology searches using the S. aureus DNA sequences allows putative

CC functions to be assigned so that protein-encoding or regulatory regions

CC of commercial, therapeutic or industrial importance can be obtained.

CC Specifically, sequences which are likely to encode antigens have been

CC identified and these polypeptides can be used in a vaccine composition

CC against S. aureus infection. The polypeptides can also be used in a kit

CC for the immunodetection of S. aureus in a sample. S. aureus is implicated

CC in numerous human diseases, including cellulitis, eyelid infections, food

CC poisoning, osteomyelitis, skin and surgical wound infections, scalded

CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the

CC DNA sequences can be used for recombinant production of the polypeptides.

CC The new DNA sequences (and their fragments) are useful as primers or

CC probes for isolating homologues of any of the 5191 S. aureus DNA sequences

CC contained on the computer readable medium

XX Sequence 936 AA;

Query Match 99.6%; Score 2690; DB 2; Length 936;

Best Local Similarity 99.6%; Pred. No. 1.7e-139;

Matches 526; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGTLLGFLSSKEADASENSVTQSDASNESKNDSSVSAAPKTDITNVSDTKTSNT 60

Db 32 VGTLLGFLSSKEADASENSVTQSDASNESKNDSSVSAAPKTDITNVSDTKTSNT 91

QY 61 NNGETSAQNPAQOQTQSSSTNATTEPTVTGEATTTTNOANTPATQSSNTNAEELV 120

Db 92 NNGETSAQNPAQOQTQSSSTNATTEPTVTGEATTTTNOANTPATQSSNTNAEELV 151

QY 121 NOTSETTNDNTVSSVNSPQNSINAENVSTQDTSATPSNNESAPQSDASNKQVY 180

Db 152 NOTSETTNDNTVSSVNSPQNSINAENVSTQDTSATPSNNESAPQSDASNKQVY 211

QY 181 NQAVNTSAPRMAFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYKLVNG 240

Db 212 NQAVNTSAPRMAFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYKLVNG 271

QY 241 FSVNPSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 300

Db 272 FSVNPSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 331

QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKTVLDVDEYKGYFYNLSIKG 360

Db 332 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKTVLDVDEYKGYFYNLSIKG 391

QY 361 TIDQIDKTNNTYRQTIYVNPNGDNVIAVLGTGNLKPNTDSNALIDQNTSIKVKVDNAA 420

Db 392 TIDQIDKTNNTYRQTIYVNPNGDNVIAVLGTGNLKPNTDSNALIDQNTSIKVKVDNAA 451

QY 421 DLSESYFVNPNPNDVTNSVNIITPFPNQYKVEFTDDQITTPYIVVNGHIDPNSKGD 480
 DB 452 DLSESYFVNPNPNDVTNSVNIITPFPNQYKVEFTDDQITTPYIVVNGHIDPNSKGD 511
 QY 481 LALRSTLYGYSNITWRSMSWDNEVAFNNGSGSGDGIDKPVVPEQDPE 528
 DB 512 LALRSTLYGYSNITWRSMSWDNEVAFNNGSGSGDGIDKPVVPEQDPE 559

RESULT 6
 AAE29262
 ID AAE29262 standard; protein; 520 AA.
 XX
 AC AAE29262;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Staphylococcus aureus Clf40 protein.
 XX
 KW Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
 KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
 KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200272600-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 28-JAN-2002; 2002WO-US002296.
 XX
 PR 26-JAN-2001; 2001US-0264072P.
 PR 12-MAR-2001; 2001US-0274611P.
 PR 18-JUN-2001; 2001US-0298413P.
 PR 30-JUL-2001; 2001US-0308116P.
 XX
 PA (INH1-) INHIBITEX INC.
 XX
 PI Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
 XX
 DR WPI; 2002-753834/82.
 DR N-PSDB; AAD46861.
 XX
 PT New anti-clumping factor A (ClfA) monoclonal antibody, useful for
 PT treating or preventing Staphylococcus aureus infection e.g. wound
 PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
 PT a human or animal.
 XX
 PS Claim 9; Page 67-69; 80pp; English.
 XX
 CC The invention relates to monoclonal antibody which binds the clumping
 CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
 CC monoclonal antibody is useful for treating or preventing S. aureus
 CC infection in a human or animal, and for inhibiting the binding of
 CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
 CC S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3
 CC protein is useful for inducing an immunological response in a human or
 CC animal. These staphylococcal infections include wound infections, sepsis,
 CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
 CC present sequence is Staphylococcus aureus Clf40 protein
 XX
 SQ Sequence 520 AA;
 Query Match 97.0%; Score 2621; DB 5; Length 520;
 Best Local Similarity 100.0%; Pred. No. 5.1e-136;
 Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 SENSVTQSDASNSKNSDSSVSAAPKTDITNYSVTYKTSNTNNGETSVQAQNETT 77
 DB 1 SENSVTQSDASNSKNSDSSVSAAPKTDITNYSVTYKTSNTNNGETSVQAQNETT 60
 QY 78 QSSSTNATTEPTVTGEATTTTNCANTPATQSSNTNABELVNOTSNETTNDINTVSS 137

DB 61 QSSSTNATTEPTVTGEATTTTNCANTPATQSSNTNABELVNOTSNETTNDINTVSS 120
 QY 138 VNSPQNSTNAENYSTTQDTSTTEATPSNNESAPQSTDSASNDKDVYVNOAVNTSAPRRAFSLA 197
 DB 121 VNSPQNSTNAENYSTTQDTSTTEATPSNNESAPQSTDSASNDKDVYVNOAVNTSAPRRAFSLA 180
 QY 198 AVAADAPAAAGTDITNQLTNVTGIDSGTTVPYHQAQYVKNLYGFSVPNSAVKGTFTKITV 257
 DB 181 AVAADAPAAAGTDITNQLTNVTGIDSGTTVPYHQAQYVKNLYGFSVPNSAVKGTFTKITV 240
 QY 258 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYI 317
 DB 241 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYI 300
 QY 318 DPNVVKKTGNVTLATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTNNTYROTII 377
 DB 301 DPNVVKKTGNVTLATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTNNTYROTII 360
 QY 378 VNSPQNSTNAENYSTTQDTSTTEATPSNNESAPQSTDSASNDKDVYVNOAVNTSAPRRAFSLA 437
 DB 361 VNSPQNSTNAENYSTTQDTSTTEATPSNNESAPQSTDSASNDKDVYVNOAVNTSAPRRAFSLA 420
 QY 438 NSVNTITFPNQYKVEFTDDQITTPYIVVNGHIDPNSKGDIALRSTLYGYSNIIWR 497
 DB 421 NSVNTITFPNQYKVEFTDDQITTPYIVVNGHIDPNSKGDIALRSTLYGYSNIIWR 480
 QY 498 SMSWDNEVAFNNGSGSGDGIDKPVVPEQDPE 528
 DB 481 SMSWDNEVAFNNGSGSGDGIDKPVVPEQDPE 511

RESULT 7
 ABUI6402
 ID ABUI6402 standard; protein; 935 AA.
 XX
 AC ABUI6402;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #1929.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029826/02.
 DR N-PSDB; ACA20272.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 44326; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression

of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
 the target prokaryotic essential genes. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

Sequence 935 AA;

Query Match 91.3%; Score 2466; DB 6; Length 935;
 Best Local Similarity 91.1%; Pred. No. 3.4e-127;
 Matches 481; Conservative 19; Mismatches 28; Indels 0; Gaps 0;

1 VGTLLGFGLLSKKEADASENSVTQSDASNESKSDSSVSAAAPKTDNVSDTKTSNT 60
 23 VGTLLGFGLLSKKEADASENSVTQSDASNESKSDSSVSAAAPKTDNVSDTKTSNT 82
 61 NNGETSVAONPAQOETQSSSTNATEETPVTEATTTTNCANTPATQSSNTAAEELV 120
 83 NNGETSVAONPAQOETQSSSTNATEETPVTEATTTTNCANTPATQSSNTAAEELV 142
 121 NQTSNETTNDNTVSSVNSPQNSNAEVSSTQDTSATPSNESAPOQSDASNKQVV 180
 143 NQTSNETTNDNTVSSVNSPQNSNAEVSSTQDTSATPSNESAPOQSDASNKQVV 202
 181 NQAVNTSAPRMRFAFLAAVAADAPAGTDITNQLNNTVVGIDSGTTPVPHQGYKLVNG 240
 203 SQAVNPSTPRMRFAFLAAVAADAPAGTDITNQLNNTVVGIDSGTTPVPHQGYKLVNG 262
 241 FSVPSAVKGDFTKLTVPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSGNVIYFTDY 300
 263 FSVPSAVKGDFTKLTVPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSGNVIYFTDY 322
 301 VNTKDDVKATLTPAYIDPENVKKTCNVLTATGISTTANKVLVDYKGYKFLNLSIKG 360
 323 VDKNENVTANITMPAYIDPENVKKTCNVLTATGISTTANKVLVDYKGYKFLNLSIKG 382
 361 TTDQIDKNTNTROTITVYNPSSGDNVIAVLGTNKLKNTDSNALIDQOQTSIKVYKVDNA 420
 383 TTDQIDKNTNTROTITVYNPSSGDNVIAVLGTNKLKNTDSNALIDQOQTSIKVYKVDNA 442
 421 DLSESYVNPENFEDVNTSVNITFNPQYKVEENTPDQITTPYIVVNGHIDNSKGD 480
 443 DLSESYVNPENFEDVNTSVNITFNPQYKVEENTPDQITTPYIVVNGHIDNSKGD 502
 481 LAIRSTLYGNGNIIWRMSWNEVAFNNGSGSGDIDKPVVPEQDDE 528
 503 LAIRSTLYGNGNIIWRMSWNEVAFNNGSGSGDIDKPVVPEQDDE 550

RESULT 8

AAU33975

ID AAU33975 standard; protein; 1021 AA.

XX AC AAU33975;

XX 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #251.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
 antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 23-OCT-2000; 2000US-0207727P.

XX 27-NOV-2000; 2000US-0242578P.

XX 22-DEC-2000; 2000US-0253625P.

XX 16-FEB-2001; 2001US-0257931P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX N-PSDB; AAS51834.

XX New polynucleotides for the identification and development of
 antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 5471; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 prokaryotic cellular proliferation, their use in identifying the genes,
 their use in the discovery of novel antibiotics, the essential genes
 themselves and the encoded proteins. The prokaryotes used are *Escherichia*
coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
Pseudomonas aeruginosa and *Enterococcus faecalis*. The invention is also
 useful for the identification of potential new targets for antibiotic
 development. The antisense nucleic acids can also be used to identify
 proteins used in proliferation, to express these proteins, and to obtain
 antibodies capable of binding to the expressed proteins. The proteins can
 be used to screen compounds in rational drug discovery programmes. The
 antisense nucleic acid sequence is also useful to screen for homologous
 nucleic acids which are required for cell proliferation in a wide variety
 of organisms. The present sequence represents an essential prokaryotic
 cellular proliferation protein. Note: The sequence data for this patent
 did not form part of the printed specification, but was obtained in
 electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1021 AA;

Query Match 90.7%; Score 2450; DB 4; Length 1021;
 Best Local Similarity 90.3%; Pred. No. 2.9e-126;
 Matches 477; Conservative 22; Mismatches 29; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKKEADASENSVTQSDASNESKSDSSVSAAAPKTDNVSDTKTSNT 60

Db 23 VGTLLGFGLLSKKEADASENSVTQSDASNESKSDSSVSAAAPKTDNVSDTKTSNT 82

QY 61 NNGETSVAONPAQOETQSSSTNATEETPVTEATTTTNCANTPATQSSNTAAEELV 120

Db 83 NNGETSVAQPAQOQETTQSASTNATTEETPTVTEATTATKQANTPATTQSSNTNABEELV 142
 QY 121 NQTSNETTNDNTVSVNSPQNSTNAENVTSTQDTSSTATPSNNSAPOSTDASNKDVV 180
 Db 143 NQTSNETASNDTNTVSVNSPQNSTNAENVTSTQDTSSTATPSNNSAPOSTDASNKDVV 202
 QY 181 NQAVNTSAPMRAPFSLAAVAADAPAAAGTDTITNLTNTVTVGIDSGTTVYVPHQAGYVKLNLYG 240
 Db 203 NQAVNTSAPMRAPFSLAAVAADAPAAAGTDTITNLTNTVTVGIDSGTTVYVPHQAGYVKLNLYG 262
 QY 241 FSPVNSAVKGDGTFKITVPKELNLNGVTSSTAKVPPIMAGDOVLANGVIDSDGNVIYTFDY 300
 Db 263 FSPVNSAVKGDGTFKITVPKELNLNGVTSSTAKVPPIMAGDOVLANGVIDSDGNVIYTFDY 322
 QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVLTATIGSTTANKVLDYKYGKFNLSIKG 360
 Db 323 VDTKENVTANITMPAYIDPENVTKTGNVLTATIGSTTANKVLDYKYGKFNLSIKG 382
 QY 361 TIDQIDKTNNTYRQTIYVNPVSGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIKVKVDNAA 420
 Db 383 TIDQIDKTNNTYRQTIYVNPVSGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIKVKVDNAN 442
 QY 421 DLSSEYFVNPENFEDVNTSVNIITPPNPQYKVFNTPDQDQITTPYIVVNGHIDPNKSGD 480
 Db 443 DLSSEYVNPSPDFEDVTNQVRISFPNANQYKVFPTDDQDQITTPYIVVNGHIDPASTGD 502
 QY 481 LALRSTLYGNSNLIWFSMSWDNEVAFNNGSGGDGIDKVPVPEQDPE 528
 Db 503 LALRSTFYGDSNFIWFSMSWDNEVAFNNGSGGDGIDKVPVPEQDPE 550

RESULT 9
 AAU36951
 ID AAU36951 standard; protein; 1021 AA.
 XX
 AC AAU36951;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Staphylococcus aureus cellular proliferation protein #1121.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI; 2001-611495/770.
 DR N-PSDB; AAS54810.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Example 3; SEQ ID NO 12544; 511pp; English.
 XX

CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1021 AA;
 Query Match 90.7%; Score 2450; DB 4; Length 1021;
 Best Local Similarity 90.3%; Pred. No. 2.9e-126;
 Matches 477; Conservative 22; Mismatches 29; Indels 0; Gaps 0;
 QY 1 VGTLLIGFGLLSKKEADASENSVTQSDASNESKNSDSSVSAAPKTDNTVSDTKTSNT 60
 Db 23 VGTLLIGFGLLSKKEADASENSVTQSDASNESKNSDSSVSAAPKTDNTVSDTKTSNT 82
 QY 61 NNGETSVAQPAQOQETTQSASTNATTEETPTVTEATTATKQANTPATTQSSNTNABEELV 120
 Db 83 NNGETSVAQPAQOQETTQSASTNATTEETPTVTEATTATKQANTPATTQSSNTNABEELV 142
 QY 121 NQTSNETTNDNTVSVNSPQNSTNAENVTSTQDTSSTATPSNNSAPOSTDASNKDVV 180
 Db 143 NQTSNETASNDTNTVSVNSPQNSTNAENVTSTQDTSSTATPSNNSAPOSTDASNKDVV 202
 QY 181 NQAVNTSAPMRAPFSLAAVAADAPAAAGTDTITNLTNTVTVGIDSGTTVYVPHQAGYVKLNLYG 240
 Db 203 NQAVNTSAPMRAPFSLAAVAADAPAAAGTDTITNLTNTVTVGIDSGTTVYVPHQAGYVKLNLYG 262
 QY 241 FSPVNSAVKGDGTFKITVPKELNLNGVTSSTAKVPPIMAGDOVLANGVIDSDGNVIYTFDY 300
 Db 263 FSPVNSAVKGDGTFKITVPKELNLNGVTSSTAKVPPIMAGDOVLANGVIDSDGNVIYTFDY 322
 QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVLTATIGSTTANKVLDYKYGKFNLSIKG 360
 Db 323 VDTKENVTANITMPAYIDPENVTKTGNVLTATIGSTTANKVLDYKYGKFNLSIKG 382
 QY 361 TIDQIDKTNNTYRQTIYVNPVSGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIKVKVDNAA 420
 Db 383 TIDQIDKTNNTYRQTIYVNPVSGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIKVKVDNAN 442
 QY 421 DLSSEYFVNPENFEDVNTSVNIITPPNPQYKVFNTPDQDQITTPYIVVNGHIDPNKSGD 480
 Db 443 DLSSEYVNPSPDFEDVTNQVRISFPNANQYKVFPTDDQDQITTPYIVVNGHIDPASTGD 502
 QY 481 LALRSTLYGNSNLIWFSMSWDNEVAFNNGSGGDGIDKVPVPEQDPE 528
 Db 503 LALRSTFYGDSNFIWFSMSWDNEVAFNNGSGGDGIDKVPVPEQDPE 550

RESULT 10
 AAU75490
 ID AAU75490 standard; protein; 496 AA.
 XX
 AC AAU75490;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE S. aureus antigenic protein associated protein #10.
 XX
 KW Antigenic protein; vaccine; SEREX; antibacterial; antiinflammatory;

dermatological; antiulcer; tuberculostatic; immunosuppressive;
septicemia; food poisoning; skin disorders; peritonitis; endocarditis;
tuberculosis; blood infection; sepsis; meningitis; pneumonia;
stomach ulcer; gonorrhea; necrotising fasciitis; impetigo;
Lyme's disease; gastro-enteritis; dysentery; shigellosis.
Staphylococcus aureus.
WO200198499-A1.
27-DEC-2001.
20-JUN-2001; 2001WO-GB002685.
20-JUN-2000; 2000GB-00014907.
(UYSH-) UNIV SHEPFIELD.
(BIOS-) BIOSYNEXUS INC.
Foster S, McDowell P, Brummell K, Clarke S;
WPI; 2002-106544/14.
Identifying antigenic polypeptides expressed by pathogenic organisms
e.g., Staphylococcus aureus during infection, by SEREX (serological
identification of antigens by recombinant expression cloning) techniques.
Disclosure; Page 72-74; 85pp; English.
The invention relates to a method for identifying antigenic polypeptides
expressed by pathogenic organisms e.g., Staphylococcus aureus during
infection, by SEREX (serological identification of antigens by
recombinant expression cloning) techniques. The method involves providing
a nucleic acid library encoding genes/partial gene sequences of
pathogenic organisms, transforming/transfecting the library into host
cells, contacting the polypeptides expressed by the genes with autologous
antisera derived from an animal infected with, or has been infected with
the pathogen and purifying the nucleic acid encoding the polypeptide or
partial polypeptide binding to the antisera. Also included are the
nucleic acids and polypeptides isolated by the method, vectors and
transformed cells expressing them, a vaccine comprising the polypeptide,
and the production of monoclonal antibodies against the polypeptides. The
protein and vaccine are useful for immunising an animal (preferably
human) against a pathogenic microbe. The proteins and antibodies are
useful for manufacturing a medicament for treating Staphylococcus aureus-
associated septicemia, food poisoning or skin disorders or
Staphylococcus epidermidis-associated septicemia, peritonitis,
endocarditis, tuberculosis, blood infections, sepsis, meningitis,
pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, impetigo,
Lyme's disease, gastro-enteritis, dysentery and shigellosis. The present
sequence is an S. aureus protein sequence associated with the antigenic
proteins of the invention. Note: The present sequence is included in the
sequence listing but is not mentioned anywhere else in the specification
Sequence 496 AA;
Query Match 89.1%; Score 2408; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.4e-124;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 VGTLLGFLGSSKEADASENSVTQSDASNSKNSDSSVSAAPKTDNTNSDTKTSNT 60
23 VGTLLGFLGSSKEADASENSVTQSDASNSKNSDSSVSAAPKTDNTNSDTKTSNT 82
61 NNGETSVAQNPAQQTQSSSTNATEETPVTEATTTTQANTPATQSSNTNABELV 120
83 NNGETSVAQNPAQQTQSSSTNATEETPVTEATTTTQANTPATQSSNTNABELV 142
121 NQTSNETTFNDNTVSSVNSPONSNTAENSTQDTSPEATPSNNESAPQSDASNKDQV 180
143 NQTSNETTFNDNTVSSVNSPONSNTAENSTQDTSPEATPSNNESAPQSDASNKDQV 202
181 NQAVNTSAPRMEAFSLAAVAADAPAGTDITNQLTNVTVGIDSGTTPVPHQAGYVKLYG 240

Db 203 NQAVNTSAPRMEAFSLAAVAADAPAGTDITNQLTNVTVGIDSGTTPVPHQAGYVKLYG 262
Qy 241 FSPVNSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAGDOVLANGVIDSGNVIYTFDY 300
Db 263 FSPVNSAVKGDFTFKITVPKELNLTGVTSTAKVPPIMAGDOVLANGVIDSGNVIYTFDY 322
Qy 301 VNTKDDVRATLTMPAYIDPENVKKTGNVLTATGIGSTTANKTTLVDYKYKGFYNLSIKG 360
Db 323 VNTKDDVRATLTMPAYIDPENVKKTGNVLTATGIGSTTANKTTLVDYKYKGFYNLSIKG 382
Qy 361 TIDQIDKNTNRYQTIYVNPNGDNVIAVLTGNLKPNTDSNALIDQONTSIKYKVDNAA 420
Db 383 TIDQIDKNTNRYQTIYVNPNGDNVIAVLTGNLKPNTDSNALIDQONTSIKYKVDNAA 442
Qy 421 DLSESYFVNPFENFEDVTNSVNTFPNPNQYKVEFNTDDQITTPYIVVNGHID 474
Db 443 DLSESYFVNPFENFEDVTNSVNTFPNPNQYKVEFNTDDQITTPYIVVNGHID 496
RESULT 11
ADA89664
ID ADA89664 standard; protein; 496 AA.
XX
ADA89664;
XX
20-NOV-2003 (first entry)
XX
XX
DE Staphylococcus aureus antigenic protein #203.
XX
antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
KW antitumor; immunostimulant; ophthalmological; pathogenic microbe;
KW bacteremia; septic shock; organ infection; skin infection;
KW bacterial basal colonisation; bacterial eye infection; septicemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.
XX
Staphylococcus aureus.
OS
WO2003011899-A2.
PN
13-FEB-2003.
PD
02-AUG-2002; 2002WO-GB003606.
PF
02-AUG-2001; 2001GB-00018925.
PR
09-JAN-2002; 2002GB-00000349.
XX
(UYSH-) UNIV SHEPFIELD.
PA
(BIOS-) BIOSYNEXUS INC.
XX
Foster S, Mond J, Clarke S, McDowell P, Brummell K;
PI
WPI; 2003-256434/25.
DR
New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
XX useful as a vaccine for immunizing humans against e.g. bacteremia, septic
XX shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhea or
XX impetigo.
PS
Claim 4; Page 161; 189pp; English.
XX
The present invention describes an antigenic protein or its part, which
is for use as a vaccine. The antigenic protein is encoded by an isolated
DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
sequences (designated dnaA and dna SE, respectively; and which encodes a
protein expressed by a pathogenic organism. Also described: (1) a vaccine
composition comprising at least one antigenic protein; (2) a method of
immunising an animal against a disease or condition caused by a
pathogenic microbe by administering the antigenic protein or the vaccine;

CC (3) an antibody or its binding part obtainable by the method above; (4)
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
CC hybridoma cell line produced by the method of (4); and (6) identifying
CC opsonic antigens expressed by a pathogenic microbe. The antigenic
CC proteins have antibacterial, immunoprotective, immunosuppressive,
CC antiinflammatory, antiulcer, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteraemia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
CC bacteria-associated food poisoning, blood infections, peritonitis,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
CC strep throat, streptococcal-associated toxic shock, necrotising
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
CC skin disorders, S. epidermidis-associated septicaemia, peritonitis or
CC endocarditis. The present sequence represents a S. aureus antigenic
CC protein sequence from the present invention.
XX
SQ Sequence 496 AA;

Query Match 89.1%; Score 2408; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.4e-124;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VTGLIGFGLSSKEADASENSVTQSDASNESKNSDSSVSAAAPKTDNTVSDTKTSNT 60
DB 23 VTGLIGFGLSSKEADASENSVTQSDASNESKNSDSSVSAAAPKTDNTVSDTKTSNT 82
QY 61 NNGETSAQNPAQOETQSSSTNATTEETPTVGEATTTTNTQANTPATTOSSTNABELV 120
DB 83 NNGETSAQNPAQOETQSSSTNATTEETPTVGEATTTTNTQANTPATTOSSTNABELV 142
QY 121 NQTSNETTFDNTVSSVNSPQNSSTNAENVSTTQDTSTEATPPSNESAPQSTDAENKDVV 180
DB 143 NQTSNETTFDNTVSSVNSPQNSSTNAENVSTTQDTSTEATPPSNESAPQSTDAENKDVV 202
QY 181 NQAVNTSAPRMRAFLAAVAADAPAACTDITNQLTNTVTGIDSGTTVYPHQAQGVKLNLYG 240
DB 203 NQAVNTSAPRMRAFLAAVAADAPAACTDITNQLTNTVTGIDSGTTVYPHQAQGVKLNLYG 262
QY 241 FSPVNSAVKGDFTKLTVPKELNGLVSTAKVPPIMAGDQVLANGVIDSGNVLTYTDDY 300
DB 263 FSPVNSAVKGDFTKLTVPKELNGLVSTAKVPPIMAGDQVLANGVIDSGNVLTYTDDY 322
QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKTVLVDYKYGKFNLSIGK 360
DB 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKTVLVDYKYGKFNLSIGK 382
QY 361 TIQIDKTNNTYRTIYVNSPGDNVIAPLVTGNLKPNTDSNALIDQNTSIKVKYVDNAA 420
DB 383 TIQIDKTNNTYRTIYVNSPGDNVIAPLVTGNLKPNTDSNALIDQNTSIKVKYVDNAA 442
QY 421 DLSESFYVNPENFEDVNSVNIITPPNPNQYKVEFNTPDQITTYIVVANGHID 474
DB 443 DLSESFYVNPENFEDVNSVNIITPPNPNQYKVEFNTPDQITTYIVVANGHID 496

RESULT 12
ADA89663
ID ADA89663 standard; protein; 496 AA.

AC ADA89663;

XX 20-NOV-2003 (first entry)

DE Staphylococcus aureus antigenic protein #202.

XX antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
KW antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
KW antiulcer; immunostimulant; ophthalmological; pathogenic microbe;
KW bacteraemia; septic shock; organ infection;

KW bacterial basal colonisation; bacterial eye infection; septicaemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.

OS Staphylococcus aureus.

PN W02003011899-A2.

PD 13-FEB-2003.

XX 02-AUG-2002; 2002WO-GH003606.

PR 02-AUG-2001; 2001GB-00018825.

PR 09-JAN-2002; 2002GB-00000349.

PA (UYSH-) UNIV SHEFFIELD.

PA (BIOS-) BIOSYNEXUS INC.

PI Foster S, Mond J, Clarke S, McDowell P, Brummel K;

DR WPI; 2003-256434/25.

PT New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
PT useful as a vaccine for immunizing humans against e.g. bacteraemia, septic
PT shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
PT impetigo.

XX Claim 4; Page 160-161; 189pp; English.

XX The present invention describes an antigenic protein or its part, which
CC is for use as a vaccine. The antigenic protein is encoded by an isolated
CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
CC sequences (designated dnaSA and dna SE, respectively; and which encodes a
CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
CC composition comprising at least one antigenic protein; (2) a method of
CC immunising an animal against a disease or condition caused by a
CC pathogenic microbe by administering the antigenic protein or the vaccine;
CC (3) an antibody or its binding part obtainable by the method above; (4)
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
CC hybridoma cell line produced by the method of (4); and (6) identifying
CC opsonic antigens expressed by a pathogenic microbe. The antigenic
CC proteins have antibacterial, neuroprotective, immunosuppressive,
CC antiinflammatory, antiulcer, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteraemia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
CC bacteria-associated food poisoning, blood infections, peritonitis,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
CC strep throat, streptococcal-associated toxic shock, necrotising
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
CC skin disorders, S. epidermidis-associated septicaemia, peritonitis or
CC endocarditis. The present sequence represents a S. aureus antigenic
CC protein sequence from the present invention.

XX SQ Sequence 496 AA;

Query Match 89.1%; Score 2408; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.4e-124;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VTGLIGFGLSSKEADASENSVTQSDASNESKNSDSSVSAAAPKTDNTVSDTKTSNT 60
DB 23 VTGLIGFGLSSKEADASENSVTQSDASNESKNSDSSVSAAAPKTDNTVSDTKTSNT 82

QY 61 NNGETSAQNPAQOETQSSSTNATTEETPTVGEATTTTNTQANTPATTOSSTNABELV 120
DB 83 NNGETSAQNPAQOETQSSSTNATTEETPTVGEATTTTNTQANTPATTOSSTNABELV 142

QY	121	NOTSNETTFNDTNTVSSVNSPQNSTNAENVSTTQDTSTEATPSNNESAPQSTDA	SKQV	180
Db	143	NOTSNETTFNDTNTVSSVNSPQNSTNAENVSTTQDTSTEATPSNNESAPQSTDA	SKQV	202
QY	181	NCVNTSAPRMAFSLAAVAADAPAAAGDTITNQLTNVTVGIDSGTTPVPHQAGYK	VLNG	240
Db	203	NCVNTSAPRMAFSLAAVAADAPAAAGDTITNQLTNVTVGIDSGTTPVPHQAGYK	VLNG	262
QY	241	FSVPNSAVKGDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVI	YTFDDY	300
Db	263	FSVPNSAVKGDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVI	YTFDDY	322
QY	301	VNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLS	IKG	360
Db	323	VNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLS	IKG	382
QY	361	TIDQIDKNTNTYRQTIYVNPVSGDNVIAPLVTGNLKPNTDSNALIDQONTSI	KVKYVDNAA	420
Db	383	TIDQIDKNTNTYRQTIYVNPVSGDNVIAPLVTGNLKPNTDSNALIDQONTSI	KVKYVDNAA	442
QY	421	DLSSSYFVNPNFEDVNTSVNMITPNNPQYKVEFNTPDQITTPYIVVNGHID	474	
Db	443	DLSSSYFVNPNFEDVNTSVNMITPNNPQYKVEFNTPDQITTPYIVVNGHID	496	

RESULT 13
AAE29263
ID AAE29263 standard; protein; 331 AA.

RESULT 13

AAE29263	
ID	AAE29263 standard; protein; 331 AA.
XX	
XX	AAE29263;
XX	
DT	27-JAN-2003 (first entry)
XX	
XX	
DE	Staphylococcus aureus Clf33 protein.
XX	
KW	Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein;
KW	immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
KW	septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.
XX	
OS	Staphylococcus aureus.
XX	
PN	WC0200272600-A2.
XX	
PD	19-SEP-2002.
XX	
PF	28-JAN-2002; 2002WO-US002296.
XX	
PR	26-JAN-2001; 2001US-0264072P.
PR	12-MAR-2001; 2001US-0274611P.
PR	18-JUN-2001; 2001US-0298413P.
PR	30-JUL-2001; 2001US-0308116P.
XX	
PA	(INH1-) INHIBITEX INC.

PI Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;

WPI: 2002-759834/82.

DR N-PSDB; AAD46862.

XX

PT New anti-clumping factor A (ClfA) monoclonal antibody, useful for
PT treating or preventing *Staphylococcus aureus* infection e.g. wound
PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
PT a human or animal.

XX
PS
Claim 9; Page 70-72; 80pp; English.

AA The invention relates to monoclonal antibody which binds the clumping
CC factor A (ClfA) protein from *Staphylococcus aureus*. The anti-ClfA
CC monoclonal antibody is useful for treating or preventing *S. aureus*
CC infection in a human or animal, and for inhibiting the binding of
CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
CC of *S. aureus* Clf40 protein, *S. aureus* Clf33 protein, or the *S. aureus* N3
CC protein is useful for inducing an immunological response in a human or

CC animal. These staphylococcal infections include wound infections, sepsis,
CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC present sequence is *Staphylococcus aureus* Clf33 protein
XX
SQ Sequence 331 AA;

Query Match	63.9%;	Score 1727;	DB 5;	Length 331;
Best Local Similarity	100.0%;	Pred. No. 3.9e-87;		
Matches 330;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	199	VAADAPAAAGTTDTNQLNTVTVGIDSGTGVVPHQAGYVKLVNGFSPNSAVKGDTEFKITVP	258	
Db	2	VAADAPAAAGTTDTNQLNTVTVGIDSGTGVVPHQAGYVKLVNGFSPNSAVKGDTEFKITVP	61	
Qy	259	KELNLNGVTSTAKVPPIMAGDQVILANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID	318	
Db	62	KELNLNGVTSTAKVPPIMAGDQVILANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID	121	
Qy	319	PENVKKTGNVTLATIGISTTANKTVLVDYEKYKGFYNLSIKGTIDQIDKTNNTYQTTIV	378	
Db	122	PENVKKTGNVTLATIGISTTANKTVLVDYEKYKGFYNLSIKGTIDQIDKTNNTYQTTIV	181	
Qy	379	NPSGDNVIAPVLTGNLKPDNTSNALIDQONTSIKVKYVDNDAADISESVFVNPENPEDVTN	438	
Db	182	NPSGDNVIAPVLTGNLKPDNTSNALIDQONTSIKVKYVDNDAADISESVFVNPENPEDVTN	241	
Qy	439	SVNITFPNPQYKVEFTPDQITTPYIVVNGHIDPNSKGDALRLSTLYGNSNIIWRS	498	
Db	242	SVNITFPNPQYKVEFTPDQITTPYIVVNGHIDPNSKGDALRLSTLYGNSNIIWRS	301	
Qy	499	MSWDNEVAFNNGSGSGDIDKPVVPEOPDE	528	
Db	302	MSWDNEVAFNNGSGSGDIDKPVVPEOPDE	331	

RESULT 14

AAW31555
ID AAW31555 standard; protein; 345 AA.
XX
XX AAW31555;
XX
XX 27-AUG-2003 (revised)
DT
DT 25-MAR-2003 (revised)
DT 21-MAY-1998 (first entry)
XX
XX
DE Fibronectin-binding MSCRAMM derivative pCF33.

XX Fibronectin; pCF33; collagen binding protein; sepsis; infection;
 KW microbial surface component recognising adhesive matrix molecule; MSCRAMM;
 KW adhesin; vaccine; immunisation; diagnosis; therapy; mastitis.
 XX
 OS *Staphylococcus aureus*.

AA	Key	Location/Qualifiers
FT	Peptide	1..12
FT		/note= "vector pOE30-derived peptide"

PN WO9743314-A2.

XXXX

PD 20-NOV-1997.

XX

PF 14-MAY-1997; 97WO-US008210.

XX
PP
16-MAY-1966
0515-0017678D

PR 16-MAY-1996; 96US-0017678P.
XX

XX PA (TEXA.) UNIV TEXAS A & M SYSTEM.

PA (UABR-) UAB RES FOUND.

XX

PI Hoeek M, Patti JM,

XX
DP 1000 000007/01

DR WPI; 1998-008801/01.
xx
PT Antibody that interacts with collagen binding domain of Staphylococcal

PT cna gene product - useful to prevent bacterial sepsis in animal infected
PT with *Staphylococcus aureus*.
XX
PS Disclosure; Page 91; 143pp; English.
XX

XX This protein comprises *Staphylococcus aureus* fibronectin-binding
CC microbial surface component recognising adhesive matrix molecule (MSCRAMM)
CC derivative pCF33, plus a vector-derived N-terminal peptide. The invention
CC relates to claimed nucleic acid sequences (see AAT93436-38) encoding S.
CC aureus collagen binding protein (CBP) epitopes M17, M31 and M55 (see
CC AAW31552-54) that confer protection against *S. aureus* infection. CBP
CC protein and antigenic epitopes are contemplated for use in the treatment
CC of pathological infections, especially to prevent bacterial adhesion to
CC collagen. The epitopes are also contemplated for use in the preparation
CC of vaccines and as carrier proteins in vaccine formulations, as well as
CC in the formulation of compositions for the prevention of *S. aureus*
CC infection. pCF33 and pCB (see AAW31556) were used to raise anti-MSCRAMM
CC polyclonal antibodies used in passive immunisation against bovine
CC mastitis. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-
CC AUG-2003 to correct OS field.)
XX
SQ Sequence 345 AA;

Query Match 63.9%; Score 1727; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 4.1e-87;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 VAAAPAAAGTDTITNQLNVTGIDSGTTPVPHQAGYKLYNFGFSPNSAVKGDFTKIVP 258
Db 14 VAAAPAAAGTDTITNQLNVTGIDSGTTPVPHQAGYKLYNFGFSPNSAVKGDFTKIVP 73
QY 259 KELNLGVTSKAPPTMAGDQVLANGVIDSGNVIYFTDYNTKDDVKATLMPAYID 318
Db 74 KELNLGVTSKAPPTMAGDQVLANGVIDSGNVIYFTDYNTKDDVKATLMPAYID 133
QY 319 PENVKKTGNVTLAGTIGSTTANKTLVDYKFKFYNLSIKGTIDQDKTNNTYRQIYV 378
Db 134 PENVKKTGNVTLAGTIGSTTANKTLVDYKFKFYNLSIKGTIDQDKTNNTYRQIYV 193
QY 379 NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 438
Db 194 NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 253
QY 439 SVNITFPNPKYKFEPTDQITTPYLVVNGHIDPNSKGLALRSTLYGNSNIWRS 498
Db 254 SVNITFPNPKYKFEPTDQITTPYLVVNGHIDPNSKGLALRSTLYGNSNIWRS 313
QY 499 MSWDNEVAFNNGSGGDIDKPVVPEQDPE 528
Db 314 MSWDNEVAFNNGSGGDIDKPVVPEQDPE 343

RESULT 15
ABU15854
ID ABU15854 standard; protein; 961 AA.
XX
AC ABU15854;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #1381.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS *Staphylococcus aureus*.
XX
FN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
FA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA19724.
PT
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 43778; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 961 AA;

Query Match 19.2%; Score 520; DB 6; Length 961;
Best Local Similarity 26.8%; Pred. No. 2.9e-20;
Matches 149; Conservative 89; Mismatches 201; Indels 118; Gaps 15;
QY 1 VGLIGFGLLSKREADASE---NSVTQSDASNESKNSDSSVSAAKPTDDTNVSDTKTS 57
Db 21 LGTMIVVGMGQEKAAASEQNNTTVESSGSATESKA-----SETQTT 63
QY 58 SNTNNGTSSVAQNPAQOQETTQSSSTNATTEPTVGTGATTTTNTQANTPATTQSSNTNAE 117
Db 64 TNNVNTIDETQSVATSTEQPSKSTQVTTTEAPTVOAPKVEETMKS----- 110
QY 118 ELVNQTSNETTFNDNTNVSSVNSPQNSTNAENYSTTQDTSTE---ATPSNESAPQSTDA 174
Db 111 -----QEDLPSEKVDKETTGTQVDIAQPSN----- 136
QY 175 SNKDVVNQAVNTSAPRMR-----AFSLAAVADAPAAAGTDTITNQLNVTGVID----SG 224
Db 137 -----VSEIKPRMKRSADVTAVSEKVEAEKATGTDVTKVKEVTESSLEGHKNDS 187

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:54:26 ; Search time 25.894 Seconds
(without alignments)
1352.277 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550

Perfect score: 2702

Sequence: 1 VGTLLGFLSSKEADASEN.....NGSGSGDIDKPVVPQDE 528

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pbp.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pbp.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pbp.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pbp.*

5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pbp.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pbp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2702	100.0	933	3	US-08-293-728-2
2	2702	100.0	933	3	US-09-421-868-2
3	2690	99.6	936	4	US-08-956-171E-5249
4	2690	99.6	936	4	US-08-781-986A-5249
5	1727	63.9	345	3	US-08-856-253-7
6	494.5	18.3	1166	4	US-09-200-650E-7
7	472	17.5	1092	4	US-09-147-405B-15
8	471.5	17.5	930	3	US-09-134-001C-5314
9	471.5	17.5	930	4	US-09-386-962C-10
10	464	17.2	918	4	US-09-200-650E-1
11	452.5	16.7	1027	4	US-08-956-171E-5254
12	452.5	16.7	1027	4	US-08-781-986A-5254
13	441.5	16.3	582	4	US-09-147-405B-13
14	441.5	16.3	593	4	US-09-147-405B-11
15	398	14.7	1315	4	US-09-200-650E-5
16	338	12.5	1155	4	US-09-710-279-1780
17	336	12.4	1742	4	US-09-386-962C-4
18	336	12.4	1742	4	US-09-386-959-4
19	327	12.1	930	4	US-09-200-650E-3
20	285.5	10.6	251	4	US-08-956-171E-5252
21	285.5	10.6	251	4	US-08-781-986A-5252
22	260.5	9.6	556	4	US-09-248-796A-22338
23	252.5	9.3	270	4	US-09-134-000C-3553
24	233.5	8.6	827	4	US-09-248-796A-17307
25	224.5	8.3	2870	4	US-09-479-467A-15
26	224.5	8.3	3178	4	US-09-479-467A-4
27	224	8.3	1112	2	US-08-714-402-2

28	224	8.3	2137	3	US-09-134-001C-4463	Sequence 4463, Ap
29	223	8.3	1161	3	US-09-327-536-2	Sequence 2, Appli
30	210.5	7.8	613	4	US-09-248-796A-17277	Sequence 17277, A
31	207	7.7	712	4	US-09-248-796A-14274	Sequence 14274, A
32	206.5	7.6	2504	4	US-09-328-352-5821	Sequence 5821, Ap
33	205	7.6	1140	4	US-09-538-092-647	Sequence 647, App
34	204	7.5	669	4	US-09-107-532A-6532	Sequence 6532, Ap
35	203	7.5	1721	3	US-08-700-651-5	Sequence 5, Appli
36	203	7.5	1721	3	US-08-928-361B-6	Sequence 6, Appli
37	203	7.5	1721	4	US-09-588-995A-6	Sequence 6, Appli
38	198.5	7.3	542	4	US-09-538-092-289	Sequence 289, App
39	197	7.3	529	4	US-09-248-796A-16703	Sequence 16703, A
40	195.5	7.2	478	4	US-09-134-000C-3709	Sequence 3709, Ap
41	194	7.2	338	4	US-09-538-092-144	Sequence 144, App
42	191.5	7.1	2315	4	US-09-543-681A-5434	Sequence 5434, Ap
43	191	7.1	1837	3	US-08-928-361B-5	Sequence 5, Appli
44	191	7.1	1837	4	US-09-588-995A-5	Sequence 5, Appli
45	190.5	7.1	983	4	US-09-538-092-1320	Sequence 1320, Ap

ALIGNMENTS

RESULT 1

US-08-293-728-2

; Sequence 2, Application US/08293728D

; Patent No. 6008341

; GENERAL INFORMATION:

; APPLICANT: Foster, Timothy J.

; APPLICANT: McDevitt, Damien L.

; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene

; FILE REFERENCE: 05344.105011

; CURRENT APPLICATION NUMBER: US/08/293,728D

; CURRENT FILING DATE: 1994-08-22

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 933

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-08-293-728-2

Query Match 100.0%; Score 2702; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 9.2e-173;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VGTLLGFLSSKEADASENSVTQSDASNESKNSDSSVSAAAPKTDITNVSDTKTSNT	60
Db	23	VGTLLGFLSSKEADASENSVTQSDASNESKNSDSSVSAAAPKTDITNVSDTKTSNT	82
Qy	61	NGETSVAQNPAAQOETTQSSSTNATTTETPVTGEATTTTNCANTPATQSSNTNABELV	120
Db	83	NGETSVAQNPAAQOETTQSSSTNATTTETPVTGEATTTTNCANTPATQSSNTNABELV	142
Qy	121	NOTSNETTNDTNTVSSVNSPQNSSTNAENVTSTQDTSTEATPSSNESAPQSTDSANKDV	180
Db	143	NOTSNETTNDTNTVSSVNSPQNSSTNAENVTSTQDTSTEATPSSNESAPQSTDSANKDV	202
Qy	181	NOAVNTSAPRMRAFLAAVAADAPAGDTINQLTNTVTGIDSGTTPVPHQGVKLNLYG	240
Db	203	NOAVNTSAPRMRAFLAAVAADAPAGDTINQLTNTVTGIDSGTTPVPHQGVKLNLYG	262
Qy	241	FSVPNSAVKGDFTKITYPKELNNGVTSTAKVPPIMAGDOVLANGVSDGNVLYTFTDY	300
Db	263	FSVPNSAVKGDFTKITYPKELNNGVTSTAKVPPIMAGDOVLANGVSDGNVLYTFTDY	322
Qy	301	VNTKDDVKALITWPAIDPENVKKTGNVLTATIGTSTANKTVLDVYEKFKFNLSIKG	360
Db	323	VNTKDDVKALITWPAIDPENVKKTGNVLTATIGTSTANKTVLDVYEKFKFNLSIKG	382
Qy	361	TIDQIDKNTNTYRTIYVNPFGDGNVIAVLITGNLKPNTDSNALIDQNTSIVKYKVDNAA	420
Db	383	TIDQIDKNTNTYRTIYVNPFGDGNVIAVLITGNLKPNTDSNALIDQNTSIVKYKVDNAA	442

QY 421 DLSESVFVNPENFEDVTSVNIITFPNPNQYKVEENTPDDQITTPYIVVNGHIDNSKGD 480
Db 443 DLSESVFVNPENFEDVTSVNIITFPNPNQYKVEENTPDDQITTPYIVVNGHIDNSKGD 502
QY 481 LALRSTLYGYNIIWRSMWNEVAFNNGSGDGDIDKPVVPEQDPE 528
Db 503 LALRSTLYGYNIIWRSMWNEVAFNNGSGDGDIDKPVVPEQDPE 550

RESULT 2

US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421,868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match 100.0%; Score 2702; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 9,2e-173;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKPTDDTNVSDTKTSNT 60
Db 23 VGTLLGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKPTDDTNVSDTKTSNT 82
QY 61 NNGETSAQNPAQOETQSSSTNATEETPVGTGEATTTTNOANTPATQSSNTAEELV 120
Db 83 NNGETSAQNPAQOETQSSSTNATEETPVGTGEATTTTNOANTPATQSSNTAEELV 142
QY 121 NOTSNETTNDNTVSSVNSPONSNAENSVTQDTSTTEATPSNNEAPQSDAENKDV 180
Db 143 NOTSNETTNDNTVSSVNSPONSNAENSVTQDTSTTEATPSNNEAPQSDAENKDV 202
QY 181 NOAVNTSAPRMRAFSLAAVAADAPAGTDTITNQLTNVTVGIDSGTTVYPHQAGYVKLYNG 240
Db 203 NOAVNTSAPRMRAFSLAAVAADAPAGTDTITNQLTNVTVGIDSGTTVYPHQAGYVKLYNG 262
QY 241 FSVPNKAVKGTDFKLTVPKELNLTGVTAKYPPIMAGDQVLANGVIDSDGNVIYFTTDY 300
Db 263 FSVPNKAVKGTDFKLTVPKELNLTGVTAKYPPIMAGDQVLANGVIDSDGNVIYFTTDY 322
QY 301 VNTKDDVKATLTPAYIDPENVKKTGNVTLATIGISTTANKVLDYDYKGYKFNLSIKG 360
Db 323 VNTKDDVKATLTPAYIDPENVKKTGNVTLATIGISTTANKVLDYDYKGYKFNLSIKG 382
QY 361 TTDQIDKNTNTRQTIYVNPNGDNVIAPIVLTGNLKNPNTDSNALIDQONTSIKVKYVDNAA 420
Db 383 TTDQIDKNTNTRQTIYVNPNGDNVIAPIVLTGNLKNPNTDSNALIDQONTSIKVKYVDNAA 442
QY 421 DLSESVFVNPENFEDVTSVNIITFPNPNQYKVEENTPDDQITTPYIVVNGHIDNSKGD 480
Db 443 DLSESVFVNPENFEDVTSVNIITFPNPNQYKVEENTPDDQITTPYIVVNGHIDNSKGD 502
QY 481 LALRSTLYGYNIIWRSMWNEVAFNNGSGDGDIDKPVVPEQDPE 528
Db 503 LALRSTLYGYNIIWRSMWNEVAFNNGSGDGDIDKPVVPEQDPE 550

RESULT 3

US-08-956-171E-5249
; Sequence 5249, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: 892-4871
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5249:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5249:
US-08-956-171E-5249

Query Match 99.6%; Score 2690; DB 4; Length 936;
Best Local Similarity 99.6%; Pred. No. 5.9e-172;
Matches 526; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKPTDDTNVSDTKTSNT 60
Db 32 VGTLLGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKPTDDTNVSDTKTSNT 91
QY 61 NNGETSAQNPAQOETQSSSTNATEETPVGTGEATTTTNOANTPATQSSNTAEELV 120
Db 92 NNGETSAQNPAQOETQSSSTNATEETPVGTGEATTTTNOANTPATQSSNTAEELV 151
QY 121 NOTSNETTNDNTVSSVNSPONSNAENSVTQDTSTTEATPSNNEAPQSDAENKDV 180
Db 152 NOTSNETTNDNTVSSVNSPONSNAENSVTQDTSTTEATPSNNEAPQSDAENKDV 211
QY 181 NOAVNTSAPRMRAFSLAAVAADAPAGTDTITNQLTNVTVGIDSGTTVYPHQAGYVKLYNG 240
Db 212 NOAVNTSAPRMRAFSLAAVAADAPAGTDTITNQLTNVTVGIDSGTTVYPHQAGYVKLYNG 271
QY 241 FSVPNKAVKGTDFKLTVPKELNLTGVTAKYPPIMAGDQVLANGVIDSDGNVIYFTTDY 300
Db 272 FSVPNKAVKGTDFKLTVPKELNLTGVTAKYPPIMAGDQVLANGVIDSDGNVIYFTTDY 331

QY	301	VNTKDDVKATLTPAYIDPENVKKTGNVTLATIGISTTANKTVLVDYKYGFYNLSIKG	360
Db	332	VNTKDDVKATLTPAYIDPENVKKTGNVTLATIGISTTANKTVLVDYKYGFYNLSIKG	391
QY	361	TIDQIDKTNNTYRQTIYVNPSGDNVIAPVLGNLKPNTDSNALIDQOQNTSIKVYKVDNAA	420
Db	392	TIDQIDKTNNTYRQTIYVNPSGDNVIAPVLGNLKPNTDSNALIDQOQNTSIKVYKVDNAA	451
QY	421	DLSESFYVNPENFEDVTVNSVNIITFPNPQYKVEFNTDDQITTPYIVVVGHIDPNSKGD	480
Db	452	DLSESFYVNPENFEDVTVNSVNIITFPNPQYKVEFNTDDQITTPYIVVVGHIDPNSKGD	511
QY	481	LAIRSTLYGYNSNIIWRSMWMDNEVAFNNGSGSGDIDKPVVPEOPDE	528
Db	512	LAIRSTLYGYNSNIIWRSMWMDNEVAFNNGSGSGDIDKPVVPEOPDE	559

RESULT 4
 US-08-781-986A-5249
 ; Sequence 5249, Application US/08781986A
 ; Patent No. 6737248
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/781,986A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Benson, Bob
 ; REGISTRATION NUMBER: 30,446
 ; REFERENCE/DOCKET NUMBER: PB248PP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 5249:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 936 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-781-986A-5249

RESULT 5
US-08-856-253-7
Sequence 7, Application US/08856253
Patent No. 6288214
GENERAL INFORMATION:
APPLICANT: Hook, Magnus
APPLICANT: Patti, Joseph M.
APPLICANT: House-Pompeo, Karen
APPLICANT: Shananam, Narayana
APPLICANT: Symersky, Jindrich
TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: U.S.
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent-Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,253
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,678
FILING DATE: 16-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: TAMK:193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
STRANDEDNESS:


```
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-13

Query Match          16.3%; Score 441.5; DB 4; Length 582;
Best Local Similarity 26.3%; Pred. No. 1e-21;
Matches 145; Conservative 99; Mismatches 227; Indels 81; Gaps 23;

QY 29 SNESKNSDSSVSAAPKTDITNV-----SDTKTSNTN---NGETSAQ 69
Db 1 SDEEKNDVINNQSIINTDNNQIIKKEETNNYDGEKSEDRTESTTNVDENEATFLOK 60

QY 70 NP-----AQQTOSSTNATTEPTVTEATTTTNOANTPATQS--SNTNAELVN- 121
Db 61 TPQDNTHLTREEVKESSSVSSNSSIDTAQQSHHTINREESVQSDNVEDSHVDFANS 120

QY 122 ---QTSNETTNDNTV-----SSVNSPQNSTNAENVSTQDTSTPATPSN---NES 167
Db 121 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN-LPINEYENKA 179

QY 168 APOSDASNKDVVQAVNTISAPRMAFSLAAVAADAPAAAGTDTITNOLTNVTVGI--DSGTT 226
Db 180 RPLSTTSQAQPSIKRVTNV-----QLAAEQGSNVNHLIKVTDQ--SITEGYDDSEGV 228

QY 227 VYPHQAQYVKLYGFSVPNSAVKGTFTKITVPKELNLTNGVTSTAKVPPIMAGD-OVLANG 285
Db 229 IKAHAENLIYDVTVEVDKVGSGDTMTVDIDKNTVPSDLTDSFIPKIDKNSGELIATG 288

QY 286 VIDS--DGNVIYTFDYVNTKDDVKATLMPAVIDPENVKKTG---NVLATGIGSTANK 341
Db 289 TYDNKNKQITFTFDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKTALESS--VNK 346

QY 342 TVLVDEYKYGKFNLSIKGTIDQIDKNTNTYQTIYVNPSPGDNVIAVLTLGNLKNPTDSN 401
Db 347 TITVBYQPNENRNTANLOSMTFNIDTKNHTVECTIYINPL--RYSAKETNNVINSNGDGE 404

QY 402 ALIDQOQNTSIKYKVDNAADLSESYFV--NPENFEDVTNSVNITFPNPNQYKVEFNTPDQ 460
Db 405 SFLIDDSIILIKYKVGDNQNLPSNRIYDYSEYEDVTNDYLAQLGNNDVNI--N 460

QY 461 ITTPYIVVNGHIDPNSKGD-----LALRSTLYGYSNIIWRSMWMDNEVAFNNGSGS 513
Db 461 IDSPYIIKVISKYDPN-KDDYTTIQOQVTMTQTINEYTGEE--FRTASYDNTIAFSTSSGQ 517

QY 514 GDGIDKPVVPEQ 525
Db 518 SQG-DLP--PEK 526

RESULT 14
US-09-147-405B-11
; Sequence 11, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:
; APPLICANT: Guss, Bengt
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Flock, Jan-Ingmar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; FILE REFERENCE: Coagulase-Negative Staphylococcus
; CURRENT APPLICATION NUMBER: US/09147405
; CURRENT FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
```

```
; PRIOR APPLICATION NUMBER: SE 9602496-3
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-11

Query Match          16.3%; Score 441.5; DB 4; Length 593;
Best Local Similarity 26.3%; Pred. No. 1.1e-21;
Matches 145; Conservative 99; Mismatches 227; Indels 81; Gaps 23;

QY 29 SNESKNSDSSVSAAPKTDITNV-----SDTKTSNTN---NGETSAQ 69
Db 8 SDEEKNDVINNQSIINTDNNQIIKKEETNNYDGEKSEDRTESTTNVDENEATFLOK 67

QY 70 NP-----AQQTOSSTNATTEPTVTEATTTTNOANTPATQS--SNTNAELVN- 121
Db 68 TPQDNTHLTREEVKESSSVSSNSSIDTAQQSHHTINREESVQSDNVEDSHVDFANS 127

QY 122 ---QTSNETTNDNTV-----SSVNSPQNSTNAENVSTQDTSTPATPSN---NES 167
Db 128 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDKISNQDELLN-LPINEYENKA 186

QY 168 APOSDASNKDVVQAVNTISAPRMAFSLAAVAADAPAAAGTDTITNOLTNVTVGI--DSGTT 226
Db 187 RPLSTTSQAQPSIKRVTNV-----QLAAEQGSNVNHLIKVTDQ--SITEGYDDSEGV 235

QY 227 VYPHQAQYVKLYGFSVPNSAVKGTFTKITVPKELNLTNGVTSTAKVPPIMAGD-OVLANG 285
Db 236 IKAHAENLIYDVTVEVDKVGSGDTMTVDIDKNTVPSDLTDSFIPKIDKNSGELIATG 295

QY 286 VIDS--DGNVIYTFDYVNTKDDVKATLMPAVIDPENVKKTG---NVLATGIGSTANK 341
Db 296 TYDNKNKQITFTFDYVDKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKTALESS--VNK 353

QY 342 TVLVDEYKYGKFNLSIKGTIDQIDKNTNTYQTIYVNPSPGDNVIAVLTLGNLKNPTDSN 401
Db 354 TITVBYQPNENRNTANLOSMTFNIDTKNHTVECTIYINPL--RYSAKETNNVINSNGDGE 411

QY 402 ALIDQOQNTSIKYKVDNAADLSESYFV--NPENFEDVTNSVNITFPNPNQYKVEFNTPDQ 460
Db 412 SFLIDDSIILIKYKVGDNQNLPSNRIYDYSEYEDVTNDYLAQLGNNDVNI--N 467

QY 461 ITTPYIVVNGHIDPNSKGD-----LALRSTLYGYSNIIWRSMWMDNEVAFNNGSGS 513
Db 468 IDSPYIIKVISKYDPN-KDDYTTIQOQVTMTQTINEYTGEE--FRTASYDNTIAFSTSSGQ 524

QY 514 GDGIDKPVVPEQ 525
Db 525 SQG-DLP--PEK 533

RESULT 15
US-09-200-650E-5
; Sequence 5, Application US/09200650E
; Patent No. 6860195
; GENERAL INFORMATION:
; APPLICANT: Fatti, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Eidehinn, Deirdre Ni
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
```

; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-5

Query Match 14.7%; Score 398; DB 4; Length 1315;
Best Local Similarity 24.5%; Pred. No. 2.5e-18;
Matches 135; Conservative 108; Mismatches 252; Indels 56; Gaps 19;

Qy	1	VGTLLIGFL--LSSKEADASENSVTQDSASNESKSDSSVSAAPKTDDT-NVSDTKTS	57
Db	37	VGTLLIGFLGNQEAKEAESTNKELNEATTSASDNQSSDKVDMQOLNQEDNTKNDNQKEMV	96
Qy	58	SNTNNGETSAQNPAQOETTQSSSTNATTEETPTVGTGATTTTNTQANTPATQSSNTNAE	117
Db	97	SSQGNETTSGNKLIEKESVQSTTGKV-----EVSTAKSDEQASPKSTNEDLNTKQ	148
Qy	118	ELVNOTSNETTFNTNTVSSV--NSPQNSTNAENVSTTQDTSTEATPSNESA-----PQ	170
Db	149	TISNCEALQPLQENKSGVNVQPTNEENKVDKATETTLNVKSDAIXNDTELVDNNSN	208
Qy	171	STDASNKDVVNQAVNTSAP-----RMRAFSLAAVAADAPAAAGTDITNQLTNVTVGIDSGT	225
Db	209	SNNENNADII--LPKSTAPKELNTRMRIAAVQPSSTEAKNVNDLITNTTLTVVDADKNN	266
Qy	226	TVYPHQAGYKLVNGFSPNSAVKGDFTKITVPKELNLNGVT-----STAKVPPIMAG	278
Db	267	KIVPAQ-DYLSLKSQITVDDKVKSGDYFTIKYSDTVQVYGLNPEDIKNIGDIKDP---NN	322
Qy	279	DQVLANGVIDSDGNVI-YTFDQVYNTKDDVKATLTMPAYIDPENVKKTGN-VTLATGIGS	336
Db	323	GETIATAKHDTANNLIITYFDYVDRNSVQMGINYIYMDADTIPVSKNDVEFNVTIGN	382
Qy	337	TTANKTVLVDYKYGKFNLSIKG---TIDQIDKTNT--YRQTIYVNPSPGDNVIAFVL	390
Db	383	TTTTKTANIQYPDYVNVNEKNSIGSAFTETVSHVGNKENPGYKQTIYVNPSENSLTNAKL	442
Qy	391	TGNLKPNTDSNAL--IDQNTSIKVKYVDNAADLSESYFVNPNFEDVTSV--NITFPN	446
Db	443	KVQAYHSSYPNNIGQINKDVTDIKIYQVPKGYTLNKGYDVNTKELTDVTNOYLQKITYGD	502
Qy	447	PNQYKVENTPDDQITTPYIVVNGHID-PNSKGDALRLSTLYGYSNIIWRSMWONEV	505
Db	503	NNSAVIDFGNAD----SAYVVMVNTKFQYTNSESPTLVQMATLSSTGN---KSVSTGNAL	555
Qy	506	AFNNGSGSGDG 516	
Db	556	GFTNNQSGGAG 566	

Search completed: November 3, 2004, 19:11:51
Job time : 27.894 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 19:03:11 ; Search time 7.73844 Seconds
(without alignments)
1078.240 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550
Sequence: 2702
1 VOTLIGFLLSKKADSEN.....NGSGSGDGIKFPVPEQDPDE 528

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 66696 seqs, 15802848 residues

Total number of hits satisfying chosen parameters: 66696

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Pending Patents AA New:*
- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
 - 2: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
 - 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
 - 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
 - 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
 - 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
 - 7: /cgn2_6/ptodata/2/paa/US11_NEW_COMB.pep.*
 - 8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199	7.4	1087	6	US-10-967-702-299
2	198	7.3	503	8	US-60-613-194-170
3	192.5	7.1	2165	6	US-10-732-923-13547
4	183	6.8	1584	6	US-10-732-923-13548
5	179.5	6.6	1430	6	US-10-797-821-36
6	168	6.2	3328	6	US-10-732-923-8311
7	165	6.1	1903	1	PCT-US04-02460-3
8	163.5	6.1	2159	6	US-10-732-923-9917
9	162.5	6.0	4315	6	US-10-739-103A-811
10	159.5	5.9	1849	6	US-10-819-275-2
11	159	5.9	3147	6	US-10-732-923-10253
12	159	5.9	3167	6	US-10-732-923-10252
13	157.5	5.8	2964	6	US-10-732-923-13552
14	156.5	5.8	585	6	US-10-868-381-58
15	155	5.7	449	6	US-10-955-952-224
16	155	5.7	449	6	US-10-157-779-224
17	155	5.7	449	6	US-10-964-241-224
18	153.5	5.7	637	6	US-10-732-923-928
19	153.5	5.7	1457	6	US-10-915-740A-1047
20	152.5	5.6	793	6	US-10-732-923-3091
21	152	5.6	1554	6	US-10-797-821-38
22	151.5	5.6	2719	6	US-10-732-923-8668
23	151	5.6	1475	6	US-10-797-821-34
24	150.5	5.6	3664	6	US-10-967-702-79
25	150	5.6	1375	6	US-10-797-821-35

26	150	5.6	1404	6	US-10-732-923-4303	Sequence 4303, App
27	147.5	5.5	1063	6	US-10-732-923-8595	Sequence 8595, App
28	146.5	5.4	585	6	US-10-868-381-57	Sequence 57, Appl
29	145.5	5.4	585	6	US-10-868-381-56	Sequence 56, Appl
30	145	5.4	1627	6	US-10-819-275-6	Sequence 6, Appl
31	145	5.4	2541	6	US-10-732-923-8761	Sequence 8761, App
32	144.5	5.3	686	6	US-10-732-923-935	Sequence 935, App
33	143.5	5.3	1026	6	US-10-478-676A-1	Sequence 1, Appl
34	143.5	5.3	1985	6	US-10-732-923-3351	Sequence 3351, App
35	143.5	5.3	6761	6	US-10-732-923-15035	Sequence 15035, A
36	142.5	5.3	1576	6	US-10-732-923-19033	Sequence 19033, A
37	142.5	5.3	2271	6	US-10-732-923-19032	Sequence 19032, A
38	142.5	5.3	2307	6	US-10-732-923-19147	Sequence 19147, A
39	142.5	5.3	2314	6	US-10-732-923-19146	Sequence 19146, A
40	142.5	5.3	4226	6	US-10-732-923-22586	Sequence 22586, A
41	142.5	5.3	4226	6	US-10-732-923-22707	Sequence 22707, A
42	142	5.3	530	8	US-60-613-207-153	Sequence 153, App
43	142	5.3	533	8	US-60-613-207-159	Sequence 159, App
44	141	5.2	419	8	US-60-613-207-151	Sequence 151, App
45	141	5.2	422	8	US-60-613-207-147	Sequence 147, App

ALIGNMENTS

RESULT 1
US-10-967-702-299
; Sequence 299, Application US/10967702
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Pauloski, Nicole
; APPLICANT: Taylor, Ian
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: GENE EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 5176
; CURRENT APPLICATION NUMBER: US/10/967,702
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/508,355
; PRIOR FILING DATE: 2003-10-03
; NUMBER OF SEQ ID NOS: 400
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 299
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-967-702-299

Query Match	7.4%	Score 199;	DB 6;	Length 1087;
Best Local Similarity	21.7%	Pred. No. 4.1e-05;		
Matches	127;	Conservative 99;	Mismatches 236;	Indels 124; Gaps 27;
Qy	11	SSKEADASENSVTQSDASN-ESKNDSSSVSAAPKTDITNVSDT-----KTSNTN	61	
Db	541	SSSQAPSSLTSTASESSSTISSNQSGYQSGPIQSTTYTSONNAQGPYQQRSTQR	600	
Qy	62	NGETSVQNAQAQBTQOS-----SSTATTETPTVTGEATTTTNTQANTPATTOSSNTNAE	117	
Db	601	RYPSSTISSSP-QXDLTQAKNGFSSVQATQQTQSVGATGSAVKSDSPST--SSIFPLN	657	
Qy	118	ELVNQTSNETTFNDNTVTVSNVSPQNSQNAENSVTTQDTTEATPNN-ESAPOSTASN	176	
Db	658	ETVSAASLLTTTQHS--SSLGGLSHSEIEPTTTTQHSSTLSQOQNTLSSTSSGRTST	715	
Qy	177	KDVVNQAVNTSA---PMRAFSLAAVAADPAAGTDITNOL---TNVTVGIDSGTTVYPH	230	
Db	716	STLTLTVESEANLHSSSTFTSTTSVTSVAPPVSVSSSLNGSSSLGLSGNSVTAS	775	
Qy	231	QAGVVKLVNFGFVSPNSAVKGDFTKITVPEKLNGLNVTSTAKVPPIMAGDOVLANGVIDSD	290	
Db	776	TRSVATTSKAPN-----LP-----FCVPLEPNPYIMAPGLLHAY	813	
Qy	291	GNVIYTFDYVNTKDDVKATLT-MPAYIDPENVKKTGNVTLATIGTGTANKTVLVYDEK	349	

Db 814 PQVYGY-----DDLQMLQTRFP--LDYYSIPFPPTPTPLTGRDGLASNPYSGDLTK 864
QY 350 YGK--FYNLSIKGTIDQID---KTNTYRQIYVNP--GDNVIA-PVLGNLKNPND 399
Db 865 FREGDASSAPATTLAQPOONQTOHTHTTQOT-FLNPAUPPGYSYLSLYTG--VPGLP 921
QY 400 SN-----ALIDQNTSIVKYVDNAADLSYFVNPENF--EDVTSNVNITFPN-----446
Db 922 STFYCGPAVFPVAPTSSKGHGVNVSNASATPFQOPSGYSGHYNTGVSTSSNTGVPDI 981
QY 447 -----PNOYKVE-----FNTDDQITTPYIVVNGHIDPNKSGDLALRSTLYGY-----490
Db 982 SGVSYSKTSQSEKQFSGTGAAGFNUFSAJGSGGPINP-----ATAAAYPPAPF 1032
QY 491 -----NSNIWRMSWNEVAFNNGSGDGDIDKPVVPEOP 526
Db 1033 MHILTPHQPHSQILHHHLQDDGQ-----TSGQBSQTSSIPQKP 1072
RESULT 2
US-60-613-194-170
; Sequence 170, Application US/60613194
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING PROTEINS ASSOCIATED WITH ABIOTIC
; TITLE OF INVENTION: STRESS RESPONSE AND PLANT CELLS AND PLANTS WITH INCREASED
; TITLE OF INVENTION: TOLERANCE TO ENVIRONMENTAL STRESS
; FILE REFERENCE: AE2004-0486
; CURRENT APPLICATION NUMBER: US/60/613,194
; CURRENT FILING DATE: 2004-09-24
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 170
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-60-613-194-170

Query Match 7.3%; Score 198; DB 8; Length 503;
Best Local Similarity 23.1%; Pred. No. 1.9e-05;
Matches 98; Conservative 76; Mismatches 181; Indels 70; Gaps 15;
QY 17 ASENVTQSDASNESKNDSSVSAAPKTD-DTNVSDTKTSNTNNGTSVAQNPAQOE 75
Db 120 ASTADSTSTATSTSTSSSTSVSSKSTKLDTKTSSGATHSSSSSTSTSTSSSE 179
QY 76 TQSSSTNATTEETPVTCGATT--TTNQANTPATQSSNTNABELVQNTSNETFNDTN 133
Db 180 TTTSSSSSSSSSTSTSTSTSSSTSTSSSTSTSSSTSTSSSTSTSTSSSTST 238
QY 134 TVSSVNSPQNSTNAENVSTTQDTSTEATPSNNESAPQSDASNKDVVNAQVNTSAPRM-- 191
Db 239 STSS-----STSTATVTSPTSSTSGTSTHYTRVVTQSVQSA-NQOASTIFT 286
QY 192 RAFSLAAVAADAPAGTDITN-----QLTNVTVGIDSGTTVPHQAGYVKL 237
Db 287 TRTSYATVTSSTSSSTSSLLNGKSSSKSLGSGAIAGVVGVVCGT-----VALLAL 340
QY 238 NYGFSV---ENSVAKGDTFKITVPEKLNNGVTSSTAKVPPIMAGDQVLANGVDSQNV 293
Db 341 ALFFVWKRRQSSQSHVDLETKQYQPSYSLGDAANPVPFSASS-----TNWHIPSRNNT 396
QY 294 IYTFDYVNTKDDVKATITMPAYIDPENVKKTGNVTATGIGSTTANKTVL--VDYKYG 351
Db 397 ALS-----KNTASTATYDLPTRAPGRDSTITGDAHNISKSHFSPVYVEPPP 445
QY 352 KFYN---LSIKGTIDQIDKNTYRQIYVNPNGDNIAPVGLNKLKPNLNDNALIDQ 407
Db 446 SIYNGORFATSPLDMME-----RQLHIVNP--DNVSSNIGSNVSGDDDDYDDAKDSN 498
QY 408 NTSIK 412
Db 499 NSSLR 503

RESULT 3

US-10-732-923-13547
; Sequence 13547, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 13547
; LENGTH: 2165
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
US-10-732-923-13547

Query Match 7.1%; Score 192.5; DB 6; Length 2165;
Best Local Similarity 18.6%; Pred. No. 0.00024;
Matches 97; Conservative 76; Mismatches 146; Indels 203; Gaps 20;
QY 18 SENVTSQSDASNESKNDSSVSAAPKTDGTDNVSDDTKTSSNTNNGTSVAQNPAQOETT 77
Db 1090 SKSSINDSSSSNNNNNN-----TTTINDSASTKGNNN---EISSPETYQKDV 1138
QY 78 QS-----SSTNA---TTEETPVTEGATTTTNOANTPATQSSNTNABELVQNT 123
Db 1139 KSKINFYBHVNISSNLNNVNTSTPTNPNVNPQSNITATTATTSTNNNNVNS 1198
QY 124 SNETTFNDTNVSSVNSPQNSTNAENVSTTQDTSTEATPSNNESAPQSDASNKDVNQA 183
Db 1199 INNNNNNN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNILSKE 1253
QY 184 VNTSAPRMRAFSLAAVAADAPAGTDITNQ--LTNTVTVGIDSGTTVYPHQAGYVKLVGF 241
Db 1254 NSTNS-----LNNLLNNTSVG-----1271
QY 242 SVPSNAVKGDTFKITVPEKLNNGVTSSTAKVPPIMAGDQVLANGVIDSGNVIYT-----296
Db 1272 KTHNRSSSG-----SDSIQPPPLPTG-----GSSHNIIFYSAFPTH 1306
QY 297 -FTDVNTKDDVKATITMPAYIDPENVKKTGNVTATGIGSTTANKTVLVDYKYGKPVN 355
Db 1307 PYTD-----IDQSTYRTT-----FGSSINNRKSLPPEYTNQVLN 1343
QY 356 LSIKGTIDQIDKNTYRQIYVNPSPGDNVIAPV-----LTGNLKPNTDS 400
Db 1344 F-----LAHSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNLENDNEN 1398
QY 401 NALIDQNTSKYKVDNAADLSYFVNPENFEDVTSNVNITFPNPNQYKVEFNTPDQ 460
Db 1399 NN-----DSFSDINDNNSVVGNDFEQ-----DDQ 1422
QY 461 I-----TTPYIVVNGHIDPNKSGDLA-----LRSTL 487
Db 1423 ILQNGISTTSTIVPTN-----DETKQLSLENSKNKVRSTV 1460

RESULT 4

US-10-732-923-13548
; Sequence 13548, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149

```
; SEQ ID NO 13548
; LENGTH: 1584
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
US-10-732-923-13548

Query Match          6.8%; Score 183; DB 6; Length 1584;
Best Local Similarity 18.4%; Pred. No. 0.00062;
Matches 121; Conservative 107; Mismatches 229; Indels 202; Gaps 31;

QY 16 DASENVTSQDSASNEKSDSSVSAAPKTDPTN--VSDTKTSSNNNGETSVQAQPAQ 73
Db 409 NSNNNNNNNNIIGKTKITTTTSTSPSSINNEDISSNNNNNNNNNNNNNNNNNN 468
QY 74 QETQSSSTNAITEETPVGEATTTTQANTPAITQSSNTNAEELVNQTSNETTENDTN 133
Db 469 NNNNNNNNNNSNNNTNNNTNNNSNNNNNNNNNSNS-----NSNSNNNNNNNN 524
QY 134 TVSSVNSPQNSTNAENVSTTQDTSSTATPSNNEAPQSTDA---SNKDVNQAV---NT 186
Db 525 NNNNNNNIYTKKPSIGTDESSTGLGNNSSGNNSSGGISGNNSSIIKORSPPHSI 584
QY 187 SAPRMAFSLAAVAADAPAGTDTNQL-----TNVTGIDSGTTPVPHQAGYVKLN---Y 239
Db 585 NGPLM-----LPPSSITNNNNIYSSYNSTTAG--SSTILP-----TLNHPIF 625
QY 240 GFSVPNS-----AVKGD-----TFKITVPKELNLN---GVSTAKVPPIMAGDQV 281
Db 626 GNTTSSNNSSSTLSVGGNNLLGRHCQSLPITASTNHTLSSSLGVSFSPSPSPKTSRK 685
QY 282 LANGVIDSGNVLYTFDY-----VNTKDDVKAILTMPAYIDPENVKKTGN 327
Db 686 IVNS--SEDGFGVQTFQDQGPSPAWRCCKSIKTKDDI--TLTI-----IKKTS 733
QY 328 VTLAGTIGSTANKTV-----LVDEYKYG-----KFYNLSIKG 360
Db 734 VAMADRPFNSSSTICYFEVYLEGHDKSGITVGLSHSTPYFKIHIGREPKSYGFSSEG 793
QY 361 TIOQIDKNTNTYQTTIYVNSPGNVIAPLVTGNLKNPTDSNALIDQON-----T 409
Db 794 EYKGGSEIGPYGPPFFFD--GDSIASSCVIG--CGINTSTRDIFFTKNGHYLGVAFSRVT 850
QY 410 SIKY-----KVDNAADLSESYF---VNPE---NFED--VT 437
Db 851 SDLYPSISPRGVVGLCVATPFGCHFRFNIEDLPGLSPSWTEALGPDQGGSGFKWAP 910
QY 438 NSVNIPTP--NPQYKVEFTPDQITTPYIVVNGHIDPNSKGDALRLSTLYGYSNI-- 494
Db 911 NDVAIWLESFNYGYRKNFR--DNNISGRHL---EGITHAMKNDLGIET--YGHREDII 963
QY 495 -----IWRSSWDN-----EYAFNNGSGSGDGI 517
Db 964 NRLNRMITQIWNDRKSPDYPKIAIDSSDKIRWPASGGSGGINISGGVYIGSSSGSDGI 1022

RESULT 5
US-10-797-821-36
; Sequence 36, Application US/10797821
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-36

Query Match          6.6%; Score 179.5; DB 6; Length 1430;
Best Local Similarity 19.3%; Pred. No. 0.00091;
Matches 106; Conservative 80; Mismatches 211; Indels 153; Gaps 21;

QY 3 TLIGFLLSKRADASENVTSQDSASNEKSDSSVSAAPKTDPTNVDTSNTNN 62
Db 31 TTLGSSVSAETEQQTSKVVYTK--SEDDKAASESSQTDAPKTKQAQTEQTOAQQANV 87
QY 63 GETSVA--QNPAAQOQETTSSTNATTEETPVTEATTTTNCQANTPAITQSSNTNAEEL 119
Db 88 ADTSTITKETSQNTTQANSDDKTV-----TNTKSEEA 122
QY 120 VNQTSNETTFNDTNTVSSVNSPQNSTNAENVSTTQDTSSTATPSNNEAPQSTDAENKOV 179
Db 123 --QTSEERT-----KQSEEAQTTASSQALTO--AKAELTKQRTAAQE-----NKNP 165
QY 180 VNQAVNTSAPRMAFSLAAVAADAPAGTDTINQLTNVTGIDSGTTPVPHQAGYVKLN 239
Db 166 VDLA---AIPNVKQI-----DGKYYIIGSDGQPKKNF 194
QY 240 GFSVPNSAVKGTFFKITVPKELNLNGVTSTAKVP-----PIMAGDOVLANGVIDSDGNVI 294
Db 195 ALFVNNKVLVFDK-----NTGALTDTSQYQFQGLTKUNNDYTPHNIQVNFENTSL 245
QY 295 YTFDVTNVTKDDVKAILTMPAYIDPENVKKTGNVTLATG-----IGSTTANKTVLVDY 347
Db 246 ETIDNV-----TADSWYRPKDLKNGKTTWATASSESDLRPLLMWWPDKQTQIAY 295
QY 348 EKY-----GKFYNLSIKGTIDQIDKNTNTYQTTIYVNSPGNVIAPLVTGNLKNPTD 399
Db 296 LNYMNOQGLGTGENY-----TADSSQESLNLAQTVOVK-----IETKISQTOQ 339
QY 400 SNALLIQQNTSIKVKYKVDNAADLSESYFVNPENFED-----VTNSVNITRPNPQYKVEFN 455
Db 340 TQWRDLINSFVQTQPNWNNSQTESDTSAGEKHDLQGGALLYSNSDKTAYAN--SDYRLNLR 398
QY 456 TPDDQITTPYIVVNGHIDPNSKG--DLALRLSTLYGYSNIIWRSMW-----DNEVAF 507
Db 399 TPTSQTKP-----KYFEDNSSGGYDFLLANDIDNSNPVVAEQNLNHLVLMYGSIVA 452
QY 508 NNGSGSGDGI 517
Db 453 NDPEANPDGV 462

RESULT 6
US-10-732-923-8311
; Sequence 8311, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8311
; LENGTH: 3328
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
```


Db 549 KSVIVVQALNNSARQTLMSDFTSYG---PTSSL---AFKPDISAPGGHIWSTQNNNGYT 602
Qy 268 S---TAKVPPIMAGDOVLANGVI--DSDGNVIYTF-----TDVNTKDDVKATLTWPA 315
Db 603 NMSGTSMASPFITAGTQALVQSQTMNDKNGAFYATYQKMSABERTPPFKITL--MNTAS:QP 660
Qy 316 YIDPENV-----KKTGNVTLATIGTSTANKTVLVDYKGYKFNLSIKG-TID-QIDK 367
Db 661 DISHNVIVPRQAGFINANATQALAKNPSTVSSNGYPGVELKFKDRTLNFQVKF 720
Qy 368 TNNYRQTIY--VNPBGDNIAPVLGTGNLKPNTDSNALIDQQNTSIKVKVDNAA--DLSES 425
Db 721 TNRKNTKALTYKLANNGK--SDVYTS---ATDSSAVLYDK-----KIDGASVKASGD 767
Qy 426 YFVNPENEDVTSNITFPN---ENQYK---VEENTPD-DQITTPYIVVNGHIDPNK 478
Db 768 IFVPANSTKELT--LTLTLPDFKENQYVEGELTFNSSDSSQLRLPYM-----GFFGDWAS 821
Qy 479 GDALRSTL 487
Db 822 SDSLPIFASL 830

RESULT 11

US-10-732-923-10253
; Sequence 10253, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10253
; LENGTH: 3147
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-732-923-10253

Query Match 5.9%; Score 159; DB 6; Length 3147;
Best Local Similarity 20.1%; Pred. No. 0.043;
Matches 123; Conservative 97; Mismatches 203; Indels 190; Gaps 30;

Qy 12 SKEADASENSVTQSDASNESKSDSSVSAPKTD--TNVSDTKTSSN-TNNGETSVA 68
Db 1931 TKQEKQORSVVVETSPSNKHIS-DESSIS---MDEIFSRQDNKSTNFKSGSIPIL 1985
Qy 69 ONPAQOETQSSSTNATTEPTVTGEATTT-----TNQA----- 103
Db 1986 VLPGEKEVASASIN--LNGVLEGGKQKSTDYGREKVKVQENSYSYSTRHMDGASIS 2043
Qy 104 -----NTPATQSSNT---NAEE-----LVNQTNETTFDNTVSSVNSPQNSTN 146
Db 2044 LDDIFNTSSTQKTETDKIDNSQEPQLSKPVLKSSI LDDLFNNTSGI-----EKT 2096
Qy 147 AENVSTTQDTSPEATPSNNEAPQSDASNDKDVNOAVN-----TSAPRMAFSLAAVAA 201
Db 2097 SKTTTTTTTTTRTET-----TDYVKRSTSLVDRFGYETATP--PAISIAAISF 2144
Qy 202 DAPAAAGTDITNQLTNVTGIDSGTTVYP---HQAGYVKLVNGYFVNSAVKGDTPKITVP 258
Db 2145 DQPS-----SSQAFFRSKHQ-----NLSSLTVP 2169
Qy 259 KELNNGVTSTAKVPPIMAGDOVLANGVIDSGNVIYFTDYNTKDDVKATLTWPA--Y 316
Db 2170 GKWNESMMSNTSTI-----SLDSDFNNSFSKNTSQV-----EPRMRKPLTLFVDNW 2217
Qy 317 IDPENVKKTGNVTATIGTSTANKTVLVDYKGYKF---YNLSIK-----GTIDQ- 364
Db 2218 ID-----NLVSEATNEATKEAPKTPKSDTLNFRSPTRISQEIKEYWADMIGDIRK 2270
Qy 365 -----IDKTNNTYRQTIYVNPBGDNIAPVLGTGNLKPNTDSNALIDQQNTSIK 413
Db 2271 NKGGHQSEEHQKHDSTWSSVVRPLETSAYTLRSSQL--NTDCKMEIDLNVFDEB 2328
Qy 414 Y-----KVDNAADLSYFVNPE-----NFEDVTN--SVNITFPNPNQYKVENTPD 458
Db 2218 ID-----NLVSEATNEATKEAPKTPKSDTLNFRSPTRISQEIKEYWADMIGDIRK 2270

Qy 365 -----IDKTNNTYRQTIYVNPBGDNIAPVLGTGNLKPNTDSNALIDQQNTSIK 413
Db 2271 NKGGHQSEEHQKHDSTWSSVVRPLETSAYTLRSSQL--NTDCKMEIDLNVFDEB 2328
Qy 414 Y-----KVDNAADLSYFVNPE-----NFEDVTN--SVNITFPNPNQYKVENTPD 458
Db 2329 LMGGKKDDEHCECSACRLTEQELBEIKKRKIELENMTSEQKIIISEPSDRKRSVDFSNPS 2388
Qy 459 ---DQITTPYIV--VUNGHDIPNSKGDALRSTLYGYSNII-----WRSMSWDN 503
Db 2389 QISLNEVSPSPVPLRIETEHAPAPSTSPPLESTRIYLLSPKVSSETVTTTHOWKQAGIPM 2448
Qy 504 EVAFNNSGSGDG 516
Db 2449 DEIFSPVSTADG 2461

RESULT 12

US-10-732-923-10252
; Sequence 10252, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10252
; LENGTH: 3167
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-732-923-10252

Query Match 5.9%; Score 159; DB 6; Length 3167;
Best Local Similarity 20.1%; Pred. No. 0.043;
Matches 123; Conservative 97; Mismatches 203; Indels 190; Gaps 30;

Qy 12 SKEADASENSVTQSDASNESKSDSSVSAPKTD--TNVSDTKTSSN-TNNGETSVA 68
Db 1931 TKQEKQORSVVVETSPSNKHIS-DESSIS---MDEIFSRQDNKSTNFKSGSIPIL 1985
Qy 69 ONPAQOETQSSSTNATTEPTVTGEATTT-----TNQA----- 103
Db 1986 VLPGEKEVASASIN--LNGVLEGGKQKSTDYGREKVKVQENSYSYSTRHMDGASIS 2043
Qy 104 -----NTPATQSSNT---NAEE-----LVNQTNETTFDNTVSSVNSPQNSTN 146
Db 2044 LDDIFNTSSTQKTETDKIDNSQEPQLSKPVLKSSI LDDLFNNTSGI-----EKT 2096
Qy 147 AENVSTTQDTSPEATPSNNEAPQSDASNDKDVNOAVN-----TSAPRMAFSLAAVAA 201
Db 2097 SKTTTTTTTTTRTET-----TDYVKRSTSLVDRFGYETATP--PAISIAAISF 2144
Qy 202 DAPAAAGTDITNQLTNVTGIDSGTTVYP---HQAGYVKLVNGYFVNSAVKGDTPKITVP 258
Db 2145 DQPS-----SSQAFFRSKHQ-----NLSSLTVP 2169
Qy 259 KELNNGVTSTAKVPPIMAGDOVLANGVIDSGNVIYFTDYNTKDDVKATLTWPA--Y 316
Db 2170 GKWNESMMSNTSTI-----SLDSDFNNSFSKNTSQV-----EPRMRKPLTLFVDNW 2217
Qy 317 IDPENVKKTGNVTATIGTSTANKTVLVDYKGYKF---YNLSIK-----GTIDQ- 364
Db 2218 ID-----NLVSEATNEATKEAPKTPKSDTLNFRSPTRISQEIKEYWADMIGDIRK 2270
Qy 365 -----IDKTNNTYRQTIYVNPBGDNIAPVLGTGNLKPNTDSNALIDQQNTSIK 413
Db 2271 NKGGHQSEEHQKHDSTWSSVVRPLETSAYTLRSSQL--NTDCKMEIDLNVFDEB 2328
Qy 414 Y-----KVDNAADLSYFVNPE-----NFEDVTN--SVNITFPNPNQYKVENTPD 458
Db 2218 ID-----NLVSEATNEATKEAPKTPKSDTLNFRSPTRISQEIKEYWADMIGDIRK 2270

```
Db 2329 LMGGKKDDEHCECSACRLTEQLEIEKKRIELENMTSEQIIISBPDRRKSVDFSNPS 2388
QY 459 -----DQITPYIV--VNGHIDPNSGDALRLTLGYNSNII-----WRMSWDN 503
Db 2389 QISLNEVFPPEVLRETETEHAPAPSTSSPPLSTRIYILSPKVSETVTTTHQWKDAGIPM 2448
QY 504 EVAFNNGSGSGDG 516
Db 2449 DEIFSPVSSTADG 2461

RESULT 13
US-10-732-923-13552
; Sequence 13552, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Egerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 13552
; LENGTH: 2964
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
US-10-732-923-13552

Query Match 5.8%; Score 157.5; DB 6; Length 2964;
Best Local Similarity 17.9%; Pred. No. 0.049;
Matches 102; Conservative 91; Mismatches 241; Indels 137; Gaps 22;

QY 1 VGTLLGGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKDDTIVNSDTKSSNT 60
Db 195 LSNIIIGFPLGNLLDP PPPPPSPNSPPIKSKSTNNLNVSNNNNNNNNNSNLSNS 254
QY 61 NNGETS---VAQNPACQETQSS--SNATTEETPVTEATTTTQACNATPATTQSSNTN 115
Db 255 GNGDESPDFQSNLVKVSRENSGNSLMLHOTSIPNNNSVVNNNNNNNNNNNNNN 314
QY 116 AELVNTQSTNETTFNDTNTVSSVNSPNTNAENVSTTQDTSTEATPSNNEAPQS--- 171
Db 315 N-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 370
QY 172 TDASNKDQVNOAVNT-----SAPRMRAFSLAAVAADAPAGTDTTNQLTNTVTG 220
Db 371 TSSGNLRLSNTENSCLCKVAIDLQNIPLYLFLSHLKI---QQLGEGKILFSAENIFC- 426
QY 221 IDSGTTVYPHQAQGVKLVNGFSVPNSAVKGD---TFKI-----TPVKELNLGV 266
Db 427 VD-----KCNF---PPNKALGEIWTNFRILFINSNNSSTIPNSTSTSI 470
QY 267 TSTA-----KVPPIMA-----GQVLANGVIDSGNVIYTFDYVNTKDDVKATLTM 313
Db 471 SSPASTQNLISLIMTLFGSYNGGSGGPTNATLSGGGS-----SSSNLNT 518
QY 314 PAYIDENVKGTGNVLTATIGIGTTANKTVLVYKYGKYNLSIKGTIDQIDKNTNTR 373
Db 519 PITTSPIHTSTNGAILNPINNLSINS-----NNNNNNNNNNNNNNNNNNNNNN-- 572
QY 374 QTIYVNPSGDNVIAPLVLTGNLKNQTDG---NALIDQNTSIKV-----YKVDNAADLSES 425
Db 573 -----INSNSNNNNINSNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 628
QY 426 YFVNPENFEDVTSNVIITFPNPQYKVEFNTPDQITTPYIVVNGHI-----DPNSKGD 480
Db 629 FKLYCKDFRCKI-----IGFQ-----INSHLVKFKDLTKCS 661
QY 481 LAIRSTLYGNSNIIWRMSGWDNEVAFNNGS 511
Db 662 VPTLDTITFAYNS-----KESSFGNTECFPDHS 688
```

```
RESULT 14
US-10-868-381-58
; Sequence 58, Application US/10868381
; GENERAL INFORMATION:
; APPLICANT: Samal, Siba K.
; TITLE OF INVENTION: Avian Pneumovirus Genes, Recombinant Avian Pneumoviruses and Metho
; TITLE OF INVENTION: Making
; FILE REFERENCE: 1797.0530002
; CURRENT APPLICATION NUMBER: US/10/868,381
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: PCT/US03/38123
; PRIOR FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US 60/430,301
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 58
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Avian pneumovirus
US-10-868-381-58

Query Match 5.8%; Score 156.5; DB 6; Length 585;
Best Local Similarity 24.7%; Pred. No. 0.008;
Matches 84; Conservative 35; Mismatches 140; Indels 81; Gaps 14;

QY 12 SKEADASENSVTQSDASNESKNSDSSVSA-----APK-----TDDTNVSDTK 55
Db 141 SKQTTEPGTSTQKHTEAPSPSSKPTTQATAOPTTAPKASTAPKNRQATTKKTETDIT 200
QY 56 TSS---NTNN-GTSAQNPACQETQSSS--TNATTEETPVTEATTTTQACNATPATT 109
Db 201 TASRARNNTNPTETATTTKATTETGKGEQHTKQEPET-TARETTPOPRRTAS 259
QY 110 QSNTNAEELVNTSNETTFNDTNTVSSVNSPNT-----NAENVSTT----- 153
Db 260 PATTKIEBAETTKRTTKN-TQTSGPRPRFSPKSTATENKRTTKTTPNSTAD 318
QY 154 -----QDTSTEATPSNNEAPQ-----STDASN-----KDVVNOAVNTS 187
Db 319 SQQTRITAEQDQQTQTRAKPTTNGAHPQTITTPHEHTDITNSTGSKPKEDKTRDPSSK 378
QY 188 APRMRAFSLAAVAADAPAG-----TDJTNQLTNTVTGIDSGTTVYPHQAQY 234
Db 379 TPTEQEDASKGTAANPGGSAAEADRRAPPATPTTGRTTESAAGTTGDSGAETRRRSAA 438
QY 235 VKLVNGFSVPNSAVKGDTFKITVPKELNLNGVTSTAKVPP 274
Db 439 DRRESTGSGTAAEAGTAQSGREAT-PKQPS-GGTAAGNTAPP 476

RESULT 15
US-10-955-952-224
; Sequence 224, Application US/10955952
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:53:31 ; Search time 109.231 Seconds
(without alignments)
2781.239 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550
Perfect score: 2702
Sequence: 1 VGTLLGFLSSKEADSEN.....NGSGSGDGDKPVVPPQDPE 528

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2702	100.0	933	2 Q53653	Q53653 staphylococ
2	2596.5	96.1	928	2 Q6GB45	Q6GB45 staphylococ
3	2596.5	96.1	928	2 Q8NXL1	Q8NXL1 staphylococ
4	2564	94.9	1029	2 Q6GJK4	Q6GJK4 staphylococ
5	2466	91.3	935	2 Q932C5	Q932C5 staphylococ
6	2466	91.3	989	2 Q99VJ4	Q99VJ4 staphylococ
7	1257	46.5	881	2 Q93KH7	Q93KH7 staphylococ
8	520	19.2	961	2 Q99RD3	Q99RD3 staphylococ
9	520	19.2	961	2 Q7A3J8	Q7A3J8 staphylococ
10	515	19.1	677	2 Q6EVM0	Q6EVM0 staphylococ
11	515	19.1	940	2 Q53682	Q53682 staphylococ
12	494.5	18.3	1166	2 Q86489	Q86489 staphylococ
13	494	18.3	943	2 Q8NUT8	Q8NUT8 staphylococ
14	494	18.3	957	2 Q6G6H4	Q6G6H4 staphylococ
15	494	18.3	1141	2 Q99W46	Q99W46 staphylococ
16	494	18.3	1141	2 Q932F7	Q932F7 staphylococ
17	492	18.2	1141	2 Q6GBS4	Q6GBS4 staphylococ
18	492	18.2	1141	2 Q8NXX5	Q8NXX5 staphylococ
19	486	18.0	965	2 Q6GDU5	Q6GDU5 staphylococ
20	481	17.8	1137	2 Q6GJA6	Q6GJA6 staphylococ
21	472	17.5	1056	2 Q8C072	Q8C072 staphylococ
22	472	17.5	1092	2 Q70022	Q70022 staphylococ
23	471.5	17.5	931	2 Q9K113	Q9K113 staphylococ
24	467	17.3	877	2 Q99R07	Q99R07 staphylococ
25	467	17.3	877	2 Q7A382	Q7A382 staphylococ
26	464	17.2	913	2 Q86476	Q86476 staphylococ
27	460.5	17.0	741	2 Q6EVM1	Q6EVM1 staphylococ
28	456.5	16.9	1171	2 Q9KWX6	Q9KWX6 staphylococ
29	452.5	16.7	1018	1 FNBA STAAU	FNBA STAAU staphylococ
30	441.5	16.3	1038	2 Q99RD2	Q99RD2 staphylococ
31	441.5	16.3	1038	2 Q7A3J7	Q7A3J7 staphylococ

ALIGNMENTS

RESULT 1									
Q53653	PRELIMINARY;		PRT;	933 AA.					
ID	Q53653								
AC	Q53653;								
DT	01-NOV-1996 (T-EMBLrel. 01, Created)								
DT	01-NOV-1996 (T-EMBLrel. 01, Last sequence update)								
DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)								
DE	Clumping factor.								
OS	Staphylococcus aureus.								
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.								
OX	NCBI_TaxID=1280;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=Newman;								
RX	MEDLINE=94224142; PubMed=8170386;								
RA	McDevitt D., Francois P., Vaudaux P., Foster T.J.;								
RT	"Molecular characterization of the clumping factor (fibrogen receptor								
RT	of Staphylococcus aureus.";								
RL	Mol. Microbiol. 11:237-248(1994).								
CC	-!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by								
CC	an amide bond (By similarity).								
DR	EMBL; Z18852; CAA79304.1; -.								
DR	PDB; 1N67; X-ray; A=202-560.								
DR	GO; GO:000986; C:cell surface; IEA.								
DR	GO; GO:0005618; C:cell wall; IEA.								
DR	GO; GO:0016020; C:membrane; IEA.								
DR	InterPro; IPR005877; Gpos_Ysirk.								
DR	InterPro; IPR001899; Gram_pos_anchor.								
DR	Pfam; PF00746; Gram_pos_anchor; 1.								
DR	Pfam; PF04650; Ysirk signal; 1.								
DR	TIGRFAMS; TIGR01167; LPXTG anchor; 1.								
DR	PROSITE; PS00847; GRAM_POS_ANCHORING; 1.								
KW	Cell wall; Peptidoglycan-anchor.								
SQ	SEQUENCE 933 AA; 97057 MW; EB51A6DE2FF759F4 CRC64;								
Query Match 100.0%; Score 2702; DB 2; Length 933;									
Best Local Similarity 100.0%; Pred.No. 1.8e-115;									
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	VGTLLGFLSSKEADSENVTQSDSAGNESKNSDSSVSAAPKTD	DDTNVSDTKTSNT	60					
Db	23	VGTLLGFLSSKEADSENVTQSDSAGNESKNSDSSVSAAPKTD	DDTNVSDTKTSNT	82					
Qy	61	NGGETSVAQPAQQTQSSSTNAATTEETPVGTAEATTINQANT	ATTCSSNTNAELV	120					
Db	83	NGGETSVAQPAQQTQSSSTNAATTEETPVGTAEATTINQANT	ATTCSSNTNAELV	142					
Qy	121	NOTSNETTFNDNTVSSVNSPQNSTNAEVSTQDTSSTATPSN	NESAPOSTDASNKDVV	180					
Db	143	NOTSNETTFNDNTVSSVNSPQNSTNAEVSTQDTSSTATPSN	NESAPOSTDASNKDVV	202					
Qy	161	NOAVNTSAPRMRAFLAAVAADAPAAAGTDITNLTNTVTGID	SGTTTYPHQAGYKLN	240					

Db 203 NOAVNTSAPRMARAFSLAAVAADAPAAAGTDIITNQLTNTVTVGIDSGTTVPYHQAGYVKLVNG 262
QY 241 FSVPSAVKGTGTFKITVPEKLNGLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
Db 263 FSVPSAVKGTGTFKITVPEKLNGLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 322
QY 301 VNTKDDVKATLTPAYIDPENVKKTGNVLTATIGTSTTANKTTLVDYKYGKFNLSIKG 360
Db 323 VNTKDDVKATLTPAYIDPENVKKTGNVLTATIGTSTTANKTTLVDYKYGKFNLSIKG 382
QY 361 TIDQIDKTNNTYRQTIYVNPISGDNVIAPLVLTGNLKNPNTDSNALIDQQNTSIKVKVDNAA 420
Db 383 TIDQIDKTNNTYRQTIYVNPISGDNVIAPLVLTGNLKNPNTDSNALIDQQNTSIKVKVDNAA 442
QY 421 DLSSEYFVNPENFEDVTNSVNIITFPNPQYKVEFNTPDQITTPYVNVVNGHIDPNSKGD 480
Db 443 DLSSEYFVNPENFEDVTNSVNIITFPNPQYKVEFNTPDQITTPYVNVVNGHIDPNSKGD 502
QY 481 LALRSTLYGYSNIIRWSMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 528
Db 503 LALRSTLYGYSNIIRWSMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 550

RESULT 2
Q6GB45
ID Q6GB45 PRELIMINARY; PRT; 928 AA.
AC Q6GB45;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Clumping factor.
GN ORFNames=SA0752;
OS Staphylococcus aureus subsp. aureus MSSA476.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSSA476;
RA Holden M.T.G., Peil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Fellwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagals K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; BX571857; CAG42526.1; -.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 928 AA; 96417 MW; FBB8570209E9B195 CRC64;

Query Match 96.1%; Score 2596.5; DB 2; Length 928;
Best Local Similarity 96.2%; Pred. No. 1.2e-110;
Matches 508; Conservative 8; Mismatches 11; Indels 1; Gaps 1;
QY 1 VGTLLGFGLLSKKEADASNSVTQSDASNSKNSDSSVSAAPKTDITNVSDTKTSNT 60
Db 23 VGTLLGFGLLSKKEADASNSVTQSDASNSKNSDSSVSAAPKTDITNVSDTKTSNT 82
QY 61 NNGETSAQNPAQOETTQSSNTNATEPTVTGEATTTTNCANTPATTQSSNTNABEVL 120

Db 83 NNGETSAQNPAQOETTQSSALTNATTEPTVTGEA-TTATNQANTPATTQSSNTNABEVL 141
QY 121 NOTSNETTNDTNTVSSVNSPQNSNAENSVSTQDTSPEATSNNESAPQSDASNKQDV 180
Db 142 NOTSNETTNDTNTVSSVNSPQNSNAENSVSTQDTSPEATSNNESAPQSDASNKQDV 201
QY 181 NOAVNTSAPRMARAFSLAAVAADAPAAAGTDIITNQLTNTVTVGIDSGTTVPYHQAGYVKLVNG 240
Db 202 NOAVNTSAPRMARAFSLAAVAADAPAAAGTDIITNQLTNTVTVGIDSGTTVPYHQAGYVKLVNG 261
QY 241 FSVPSAVKGTGTFKITVPEKLNGLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
Db 262 FSVPSAVKGTGTFKITVPEKLNGLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 321
QY 301 VNTKDDVKATLTPAYIDPENVKKTGNVLTATIGTSTTANKTTLVDYKYGKFNLSIKG 360
Db 322 VNTKDDVKATLTPAYIDPENVKKTGNVLTATIGTSTTANKTTLVDYKYGKFNLSIKG 381
QY 361 TIDQIDKTNNTYRQTIYVNPISGDNVIAPLVLTGNLKNPNTDSNALIDQQNTSIKVKVDNAA 420
Db 382 TIDQIDKTNNTYRQTIYVNPISGDNVIAPLVLTGNLKNPNTDSNALIDQQNTSIKVKVDNAA 441
QY 421 DLSSEYFVNPENFEDVTNSVNIITFPNPQYKVEFNTPDQITTPYVNVVNGHIDPNSKGD 480
Db 442 DLSSEYFVNPENFEDVTNSVNIITFPNPQYKVEFNTPDQITTPYVNVVNGHIDPNSKGD 501
QY 481 LALRSTLYGYSNIIRWSMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 528
Db 502 LALRSTLYGYSNIIRWSMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 549

RESULT 3
Q8NXJ1
ID Q8NXJ1 PRELIMINARY; PRT; 946 AA.
AC Q8NXJ1;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein.
GN Name=clfa; OrderedLocusNames=MW0764;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yanamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AP004824; BAB94629.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 946 AA; 98237 MW; EFPB838793201173 CRC64;

Query Match 96.1%; Score 2596.5; DB 2; Length 946;
Best Local Similarity 96.2%; Pred. No. 1.2e-110;
Matches 508; Conservative 8; Mismatches 11; Indels 1; Gaps 1;

QY 1 VGLTIGFGLLSKEADASENSVTQSDASNESKSDSSVSAAPKTDITDINVSDDTKSSNT 60
DB 23 VGLTIGFGLLSKEADASENSVTQSDASNESKSDSSVSAAPKTDITDINVSDDTKSSNT 82
QY 61 NNGETSAVQAQOQETQSSNATTEETPVTEGATTTTNTQANTPATTOSSNTNAEELV 120
DB 83 NNGETSAVQAQOQETQSSNATTEETPVTEGATTTTNTQANTPATTOSSNTNAEELV 141
QY 121 NQTSNETTNDNTVSSVNSPQNSNAENVTQDTSSTATSNNSAPQSDTASNDKVV 180
DB 142 NQTSNETTNDNTVSSVNSPQNSNAENVTQDTSSTATSNNSAPQSDTASNDKVV 201
QY 181 NOAVNTSAPRMRAFSALAAVAADAPAAAGTDTITNQLTNVTVGIDSGTTVYVPHQAGYVKLYNG 240
DB 202 NOAVNTSAPRMRAFSALAAVAADAPAAAGTDTITNQLTNVTVGIDSGTTVYVPHQAGYVKLYNG 261
QY 241 FSVPSAVKGDFTKFTVPKELNNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDY 300
DB 262 FSVPSAVKGDFTKFTVPKELNNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDY 321
QY 301 VNTKDDVKATLMPAYIDPENVKKTGNVTLAGTIGSTTANKTVLVDYKGYKFNLSIKG 360
DB 322 VNTKDDVKATLMPAYIDPENVKKTGNVTLAGTIGSTTANKTVLVDYKGYKFNLSIKG 381
QY 361 TIDQIDKTNNTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKYKVDNAA 420
DB 382 TIDQIDKTNNTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKYKVDNAA 441
QY 421 DLSESYFVNPENPEDVTNSVNIITFPNPQYKVEFTPDQOITTPYIYVWNGHIDPNKSGD 480
DB 442 DLSESYFVNPENPEDVTNSVNIITFPNPQYKVEFTPDQOITTPYIYVWNGHIDPNKSGD 501
QY 481 LALRSTLYGYSNIIWRSMWDNEVAFNNGSGSGDGIKXPVPEQDPE 528
DB 502 LALRSTLYGYSNIIWRSMWDNEVAFNNGSGSGDGIKXPVPEQDPE 549

RESULT 4

Q6GIK4 PRELIMINARY; PRT; 1029 AA.
AC Q6GIK4; (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE Clumping factor.
GN Name=clfa; ORFNames=SR0842;
OS Staphylococcus aureus subsp. aureus MRSA252.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282458;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MRSA252;
RA Holden M.T.G., Feil R.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holtroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mays R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR ENBL; BX571856; CAG39851.1; -;
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK_signal; 1.

DR PROSITE; P50847; GRAM_POS_ANCHORING; 1.
KW Cell wall, Peptidoglycan-anchor.
SQ SEQUENCE 1029 AA; 106715 MW; 025B13C6A4C4F020 CRC64;
Query Match 94.9%; Score 2564; DB 2; Length 1029;
Best Local Similarity 94.1%; Pred. No. 4.1e-109;
Matches 497; Conservative 17; Mismatches 14; Indels 0; Gaps 0;
QY 1 VGLTIGFGLLSKEADASENSVTQSDASNESKSDSSVSAAPKTDITDINVSDDTKSSNT 60
DB 23 VGLTIGFGLLSKEADASENSVTQSDASNESKSDSSVSAAPKTDITDINVSDDTKSSNT 82
QY 61 NNGETSAVQAQOQETQSSNATTEETPVTEGATTTTNTQANTPATTOSSNTNAEELV 120
DB 83 NNGETSAVQAQOQETQSSNATTEETPVTEGATTTTNTQANTPATTOSSNTNAEELV 142
QY 121 NQTSNETTNDNTVSSVNSPQNSNAENVTQDTSSTATSNNSAPQSDTASNDKVV 180
DB 142 NQTSNETTNDNTVSSVNSPQNSNAENVTQDTSSTATSNNSAPQSDTASNDKVV 202
QY 181 NOAVNTSAPRMRAFSALAAVAADAPAAAGTDTITNQLTNVTVGIDSGTTVYVPHQAGYVKLYNG 240
DB 203 NOAVNTSAPRMRAFSALAAVAADAPAAAGTDTITNQLTNVTVGIDSGTTVYVPHQAGYVKLYNG 262
QY 241 FSVPSAVKGDFTKFTVPKELNNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDY 300
DB 263 FSVPSAVKGDFTKFTVPKELNNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDY 322
QY 301 VNTKDDVKATLMPAYIDPENVKKTGNVTLAGTIGSTTANKTVLVDYKGYKFNLSIKG 360
DB 323 VNTKDDVKATLMPAYIDPENVKKTGNVTLAGTIGSTTANKTVLVDYKGYKFNLSIKG 382
QY 361 TIDQIDKTNNTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKYKVDNAA 420
DB 383 TIDQIDKTNNTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKYKVDNAA 442
QY 421 DLSESYFVNPENPEDVTNSVNIITFPNPQYKVEFTPDQOITTPYIYVWNGHIDPNKSGD 480
DB 443 DLSESYFVNPENPEDVTNSVNIITFPNPQYKVEFTPDQOITTPYIYVWNGHIDPNKSGD 502
QY 481 LALRSTLYGYSNIIWRSMWDNEVAFNNGSGSGDGIKXPVPEQDPE 528
DB 503 LALRSTLYGYSNIIWRSMWDNEVAFNNGSGSGDGIKXPVPEQDPE 550

RESULT 5

Q932C5 PRELIMINARY; PRT; 935 AA.
AC Q932C5; (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein.
GN Names=fnb; OrderedLocustNames=SAV0811;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-O., Ito T.,
RA Kananori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizukami-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshida K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).

DR EMBL; AP003360; BAB56973.1; --
DR HSSP; Q53653; 1N67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR Pfam; PF00746; Gram_pos_anchor.
DR Pfam; PF04650; Y5IRK signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 935 AA; 96949 MW; DCSA2D92CE3BA91C CRC64;

Query Match
Best Local Similarity 91.1%; Score 2466; DB 2; Length 935;
Matches 481; Conservative 19; Mismatches 28; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKKEADASENSVTQSDASNESKNDSSVSAAPKTDNTVSDTKTSNT 60
Db 23 VGTLLGFGLLSKKEADASENSVTQSDASNESKNDSSVSAAPKTDNTVSDTKTSNT 82
QY 61 NNGETSAQNPAQOETTQSSSTNATTEPTVGTGEATTTTNCANTPATQSSNTNAEELV 120
Db 83 NNGETSAQNPAQOETTQSSSTNATTEPTVGTGEATTTTNCANTPATQSSNTNAEELV 142
QY 121 NQTSNETTFNDTNTVSSVNSPQNSTNAENVSTQDTSNNEAPQSDASNKDGV 180
Db 143 NQTSNETTNDTNTVSSVNSPQNSTNAENVSTQDTSNNEAPQSDASNKDGV 202
QY 181 NQAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQLNVTGIDSGTIVYPHQAQYVKLYNG 240
Db 203 SQAVNPSTPRMRAFSLAAVAADAPAAAGTDITNQLTDVKVTDISGTTVYPHQAQYVKLYNG 262
QY 241 FSVNSAVKGDTEKFTVPEKELNAGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
Db 263 FSVNSAVKGDTEKFTVPEKELNAGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 322
QY 301 VNTKDDVKATLTPAYIDPENVKKTGNVLTATGISTTANKTTLVDYKYGKFNLSIKG 360
Db 323 VDNKENVNTANITMPAYIDPENVKKTGNVLTGIGTNTASKTTLVDYKYGKFNLSIKG 382
QY 361 TIDQIDKTNNTYRTIYVNPAGDVIAPVLTGNLKNPNTDSNALIDQOQNTSIKVKVDNAA 420
Db 383 TIDQIDKTNNTYRTIYVNPAGDVIAPVLTGNLKNPNTKSNALIDAKNTDIKVVVDNAN 442
QY 421 DLSSEYFVNPENFEDVTNSVNIITPENQYKVEPTDDQITTPYIVVNGHIDPNSKGD 480
Db 443 DLSSEYFVNPENFEDVTNQVRISFPNANQYKVEPTDDQITTPYIVVNGHIDPASTGD 502

RESULT 6
Q99VJ4
ID Q99VJ4 PRELIMINARY; PRT; 989 AA.
AC Q99VJ4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DE Fibrinogen-binding protein A, clumping factor.
GN Name=ClfA; Ordered locus Names=SA0742;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iino J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,

RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240 (2001).
CC - SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AP003131; BAB41975.1; --
DR PIR; D89852; D89852.
DR HSSP; Q53653; 1N67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 989 AA; 102407 MW; DA6B807539623467 CRC64;

Query Match
Best Local Similarity 91.3%; Score 2466; DB 2; Length 989;
Matches 481; Conservative 19; Mismatches 28; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKKEADASENSVTQSDASNESKNDSSVSAAPKTDNTVSDTKTSNT 60
Db 23 VGTLLGFGLLSKKEADASENSVTQSDASNESKNDSSVSAAPKTDNTVSDTKTSNT 82
QY 61 NNGETSAQNPAQOETTQSSSTNATTEPTVGTGEATTTTNCANTPATQSSNTNAEELV 120
Db 83 NNGETSAQNPAQOETTQSSSTNATTEPTVGTGEATTTTNCANTPATQSSNTNAEELV 142
QY 121 NQTSNETTFNDTNTVSSVNSPQNSTNAENVSTQDTSNNEAPQSDASNKDGV 180
Db 143 NQTSNETTNDTNTVSSVNSPQNSTNAENVSTQDTSNNEAPQSDASNKDGV 202
QY 181 NQAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQLNVTGIDSGTIVYPHQAQYVKLYNG 240
Db 203 SQAVNPSTPRMRAFSLAAVAADAPAAAGTDITNQLTDVKVTDISGTTVYPHQAQYVKLYNG 262
QY 241 FSVNSAVKGDTEKFTVPEKELNAGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
Db 263 FSVNSAVKGDTEKFTVPEKELNAGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 322
QY 301 VNTKDDVKATLTPAYIDPENVKKTGNVLTATGISTTANKTTLVDYKYGKFNLSIKG 360
Db 323 VDNKENVNTANITMPAYIDPENVKKTGNVLTGIGTNTASKTTLVDYKYGKFNLSIKG 382
QY 361 TIDQIDKTNNTYRTIYVNPAGDVIAPVLTGNLKNPNTDSNALIDQOQNTSIKVKVDNAA 420
Db 383 TIDQIDKTNNTYRTIYVNPAGDVIAPVLTGNLKNPNTKSNALIDAKNTDIKVVVDNAN 442
QY 421 DLSSEYFVNPENFEDVTNSVNIITPENQYKVEPTDDQITTPYIVVNGHIDPNSKGD 480
Db 443 DLSSEYFVNPENFEDVTNQVRISFPNANQYKVEPTDDQITTPYIVVNGHIDPASTGD 502

RESULT 7
Q93MH7
ID Q93MH7 PRELIMINARY; PRT; 881 AA.
AC Q93MH7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Receptin Fbl precursor.

GN	Name=fb1;
OS	Staphylococcus lugdunensis.
OC	Bacteria; Firmicutes; Bacillales;
OX	NCBI Taxid=28035;

[1] [_GenBank/EMBL/CC](#)
 RN SEQUENCE FROM N.A.
 RP STRAIN=2342;
 RC
 RA Nilsson M., Ahlen J., Frykberg L., Guss B.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC - - SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond.(By similarity)
 CC

```

DR EMEL; AF404823; AAK95649.1; -.
DR HSP; Q53653; In67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMS; TIGR01167; LPXTC_anchor; 1.
DR TIGRFAMS; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PSS0847; GRAM_POS_ANCHORING; 1.
DR KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 39 Potential.
FT CHAIN 40 881
SQ SEQUENCE 881 AA; 94250 MW; D4296C4959CA4F19B CRC64;

```

Query Match	46.5%	Score 1257;	DB 2;	Length 881;
Best Local Similarity	47.2%	Pred. No. 1.6e-49;		
Matches 259;	Conservative	82;	Mismatches 150;	
			Indels 58;	Gaps 10;

QY	1	VGLILGGLLSSKEADASE-----NSVQTQDSASNESKGNDSSSVSAAAP	44
		: : : : : : : : :	
Db	23	VGSLIGLMLMTKEAEAEVERNLKQOI QHNNDATGTDQDNNYNNEIQEATFNQKQ	82
		: : : : : : : : :	
QY	45	KTDDTNV-SDTKTSNTNNGETSAQNPAAQQTQSSSTNATTEETPVTGEATTTT--TN	101
		: : : : : : : : :	
Db	83	ITQSDNVNSAQAINELSDSHRV--NKATEALDNNSTLNTSIDVSPATKODTTTSMOTT	140
		: : : : : : : : :	

QY	102	QANTPATQSSNTNAEELVNVQTSNETTFTNTVSSVNSPQSTNAENVSTQDTSTPEAT	161
		: : : : : : : : :	
Db	141	QENNDAITQTKVYKQDGNVNLQVATNDN---QSSNQPRNS-----HLNST-VT	198
QY	162	PSNNEGAPO--STDASKDVVNVQAVNTSAPEPMRAFSLAAVAADAPAAAGTQITINQLTNVTV	219
Db	189	YNNHQVRELAKVEATND--NNVTQTS-----DISNKLNVTA	225

220	GIDSGTTVPHQAGYVKLVNGFSPNSAVKGDFTKIVRKELNNGVTSTAKVPPTMAGD	279
QY		
226	TISPAADTIYHKAEYVNLNTRFOAPDVGDSIKITIPQALNNGVYTAQAKPNITMAGD	285
Db		

280	QVLANGVPSDGNVITFTFDVNTKDDVKATLWMPAVIDPENVKKTGNVTLAGTGSTTA	339
QY		
285	QILATGDIIEGNLIYTFDVTNNKNTIGTISPIGVIDPNKNTHTGKNTFSTIGTTA	345
Db		

20	QY	340	NKTVLVDYKVGKFNLSIKGTIDQDKTNTYRQTIVYVPSGDNVIAPVLGNLKPNTD	399
200	QY	346	KKTVTDYKVGKFNLSIKGTIDQDKTNTYRQTIVYVPSGDNVIAPVLGNLKPNTD	405
200	QY	346	KKTVTDYKVGKFNLSIKGTIDQDKTNTYRQTIVYVPSGDNVIAPVLGNLKPNTD	405

346 KALVIVDIEKIGEFNLSINGLIDLDUKNNITKQVYVNPSSDTVDDPYLRGSGIPGTT 405

QY 400 SNALIDQONTISIKVYKVDNAADISESYFVNPEFEDVTSVNTFFNPQYKVEFVTPDD 459

Db 406 SNVIDEONTISIKVYKVEKZAHITQSYVNPSSKVENVTNVKTEPEKGYGVNWEPTDD 458

LD	406	SNVILDEQNTISIAVYKVKAKHUTDSYVDPFSNYEDVTSVKITPFBKGIYQINFNTEDD	465
QY	460	QITPTPVVVVNGHIDPNSKGDALRLSTLYGYNIIWRSWSDNEVAFNNGSSGGIDK	519
D6	466	QNSPVPVATNGVHDNSGNIYVISTLYCYGNSGEMVGMWNNVEVYVCSGNGCTVY	505

466 QINSFVWINGHVDNNSNGLYRLSTLYGDSNFTRVSNMWDNEVEYHAGSGNGDIDK 525

db 526 PVIPDQPGD 534

RESULT 8

Q99RD3	PRELIMINARY;	PRT;	961 AA.
AC	Q99RD3;		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Fibronectin-binding protein homolog.		
GN	Name=fnbB; OrderedLocusNames=SAV2502;		
OS	Staphylococcus aureus (strain Mu50 / ATCC 700699).		
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
OX	NCBI_TaxID=158878;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Mu50 / ATCC 700699;		
RX	MEDLINE=21111952; PubMed=11418146;		
RA	Kuroda M., Ohita T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,		
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,		
RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshiyama A.,		
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,		
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,		
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,		
RA	Hattori M., Ogasawara N., Hayashi H., Hiranatsu K.,		
RT	"whole genome sequencing of methicillin-resistant Staphylococcus		
RL	aureus."		
RL	Lancet 357:1225-1240(2001).		

CC CC
CC CC

Lancel 357:1225-1240(2001).
--!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (By similarity).

DR EMBL; AP003365; BAB58664.1; -.
DR PIR; G90053; G90053.

DR HSSP; Q53653; 1N67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO:0009986; C:cell surface; IEA.

DR	GO: 0009275;	C: cell wall (sensu Gram-positive Bacteria); IEA.
DR	GO: 0016020;	C: membrane; IEA.
DR	GO: 0009275;	C: cell wall (sensu Gram-positive Bacteria); IEA.
DR	GO: 0016020;	C: membrane; IEA.

```
DR GO:0007155; P:cell adhesion; IEA.  
DR InterPro: IPR004237; Fm_bind.  
DR GO:0007155; P:cell adhesion; IEA.  
DR InterPro: IPR004237; Fm_bind.
```

```

DR      InterPro; IPR005877; Gpos_Y5IRK.
DR      InterPro; IPR001899; Gram_pos_anchor.
DR      InterPro; IPR000000; Gram_pos_anchor.

```

DR pfam; PF02986; Ff_bind; 1.
DR pfam; PF00746; Gram_pos_anchor; 1.
DR pfam; PF04550; vsmv; 1.
DR pfam; PF00746; Gram_pos_anchor; 1.
DR pfam; PF02986; Ff_bind; 1.

```
DR pfam; pf04650; YSIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; VSPV_signal; 1.
DR TIGRFAMs; TIGR01169; VSPV_signal; 1.
```

DR DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
DR DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
KW KW Cell wall; Complete proteome. Pentideolycan-anchor

RAW	Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ	SEQUENCE 961 AA; 106010 MW; 364940F884EA4101 CRC64;

Query Match	19.2%;	Score 520;	DB 2;	Length 961;
Best Local Similarity	26.8%;	Pred. No. 8.1e-16;		

Local Similarity 20.8%; Fied. NO. 8.1E-16;
Matches 149; Conservative 89; Mismatches 201; Indels 118;

Qy **1** VGTLLGFGLLSKEADASE---NSVTQSDSASNESKSNDSSSVSAAPKTDDTNVS
 :|||:|: ||| ||| ||| :||| :||| :||| :||| :||| :||| :

Db 21 LGTIVVGMGQKEKAAASEQNNTTVEESGSSATESKA-----S

QY 58 SNTNGETSVAQNPAQQETTQSSSTNATTBETPVTGEATTTTNTQANTPATQSS

Db 64 TNNVNIDETQSYSATSTEQPSKSTQVTTEAPTIVQAPKVEEMKS

[illegible]

DB	111	-----QEDLPSEKVDKETTTGTQVDAQPNS-----
OV	175	SNKWYWNQAVNTSADBBMB-----APSIKAUVAADADAACGRTTNOATTAATTTCCTD

```

QY      175  SNKDVWNQAVNTSAPRMR-----AFSLAAVAADAPAAAGTDITNQLTNVTGID
          :   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      137  -----VSEIKPBMBKSADVTAUVEKEVAFEAKATCTGTNTNKVETTESSEI

```

225 TTYVPHOAGYVKI NYGESVDPNSAVKKGDTFKI TVPKELNINVTSTAKVPDIMA-
 13 / -----VSEIKPRMRSADVIAVSEKEVAEEAKATGTDVINKVENTESLEG-
 DD -----

```

188 NIWPHNAORVTLKYKWKFEFGIKAGDYDFETLSDNVETHGISTIRKVPKESKST
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
225 IIVIPHQAGIVKLNLYG:SVENSAVKGDIFKITVPRELNLNGWVSTAKVVPPIMAG
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  QY      QY      QY      QY      QY      QY      QY      QY      QY
  Db      Db      Db      Db      Db      Db      Db      Db      Db

```

DZ		188	N I V N F H N A Q V I L A I R W K P E G E I K A G D I F D F I L S D U N V E T H G I S I L K K V P E I K R S S
DZ		283	A N G V I D S D G N V I Y T F D Y V N T K D P V K A T L T M P A Y I D P E N V K X T G N V T L A T G I G S T
Qy			

263 ANGVLESDGNVLIIFILINLVNDVNALLIMFAIDFENVANIGNVLIAIGS

```
Db 248 ANGVINERTIRYTFDYINNKKOLTAELNLFIDPTTVTKQSGQKVEVTLGQNKVSKE 307
QY 343 VLVDY-----EKYGFYNLSIKGTIDQIDKTNNTYRQIYVNPSCDGNVIAPIVLTGNLKEN 397
Db 308 FDIKYLGVKDRMG-----VTVNGRIDTLNKEGKFSHFAYVKPNQSLTSVTVTGQVTS 363
QY 398 TDSNALIDQNTSIRKVVYKVDNAADLSSEYFV---NPNFEDVTNSVNIFFNPNOYKVEF 454
Db 364 YKQSA---NNPTVKVYKHIGSDELAESVYAKLDTSKFEDVTEKVNLSYTSNGGYTLNL 419
QY 455 NTPDQIITPYIVVNGHIDPNSKGDALRSTLYGYNLSNIWR-----SMSWDNEVA 506
Db 420 GDLDN--SKDYVYKVEGEYDQNAK--DLNFRHLSGYHKYYPYYPYPPVQLTWNNGVA 476
QY 507 FNNGSGSGDGIDKPVP 523
Db 477 FYSNNAKGDKKPNDP 493

RESULT 9
Q7A3J8 PRELIMINARY; PRT; 961 AA.
ID Q7A3J8
AC Q7A3J8;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE FndB protein
GN Name=fnbB; OrderedLocusNames=SA2290;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RA "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AF003137; BAB43593.1; -.
DR InterPro; IPR004237; Fb_bind.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF02986; Fb_bind; 1.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 961 AA; 106010 MW; 364940F894EA4101 CRC64;

Query Match 19.2%; Score 520; DB 2; Length 961;
Best Local Similarity 26.8%; Pred. No. 8.1e-16;
Matches 149; Conservative 89; Mismatches 201; Indels 118; Gaps 15;

QY 1 VGTLLIGFLSSKEADASE---NSVTQSDSASNESKSSDSSVSAAPKTDITNVSDTKTS 57
Db 21 LGTMIVGMGQKEAAASEQNNTTVESGSSATESKA-----SETQTT 63
QY 58 SNTNNGETSAQNPAQOQETQSSSTNATTEETPVTEATTTTNOANTPATTQSSNTNAE 117
Db 64 TNNVNTIDETQSYATSSTEQPSKSTQVTTTEAPTQVQAPKVEETMSK-----110
QY 118 ELVNQTSNETTNDNTVSSVNSPQNSNAENVSTTQDTSSTE---ATPSNESAPOSTDA 174
Db 133 -----VSEIKPRMKESTDVTAAVEKEVVEETKATGCTVTK-----VEVEGSEIV 179
```

```
Db 111 -----QEDLPSEKADKETTGTVQVDTAQPNS-----136
QY 175 SNKDVVNQAVNTSAPRMR-----AFSLAAVAADAPAAAGTDITNLTNTVTGID-----SG 224
Db 137 -----VSEIKPRMKESADVTAVSEKEVAEAKATGTDVTKVEVTESSLEGHKDS 187
QY 225 TTVPHQAGYVKLIANGFSVPNSAVKGDTFKTIIVPKELNLTNGVTSTAKUPPIMAG--DQVL 282
Db 188 NVNPHNAQRVTLKYKWKFGEGIKAGDYFDTLSDNVETHGISTLRKVPKISTSEDKVM 247
QY 283 ANGVIDSNGVIVYFTDYVNTKDVKA TLTPAYIDPENVKKTGNVLTATGISTTANKT 342
Db 248 ANGVINERTIRYTFDYINNKKOLTAELNLFIDPTTVTKQSGQKVEVTLGQNKVSKE 307
QY 343 VLVDY-----EKYGFYNLSIKGTIDQIDKTNNTYRQIYVNPSCDGNVIAPIVLTGNLKEN 397
Db 308 FDIKYLGVKDRMG-----VTVNGRIDTLNKEGKFSHFAYVKPNQSLTSVTVTGQVTS 363
QY 398 TDSNALIDQNTSIRKVVYKVDNAADLSSEYFV---NPNFEDVTNSVNIFFNPNOYKVEF 454
Db 364 YKQSA---NNPTVKVYKHIGSDELAESVYAKLDTSKFEDVTEKVNLSYTSNGGYTLNL 419
QY 455 NTPDQIITPYIVVNGHIDPNSKGDALRSTLYGYNLSNIWR-----SMSWDNEVA 506
Db 420 GDLDN--SKDYVYKVEGEYDQNAK--DLNFRHLSGYHKYYPYYPYPPVQLTWNNGVA 476
QY 507 FNNGSGSGDGIDKPVP 523
Db 477 FYSNNAKGDKKPNDP 493

RESULT 10
Q6EVM0 PRELIMINARY; PRT; 677 AA.
ID Q6EVM0
AC Q6EVM0;
DT 01-OCT-2004 (TReMBLrel. 28, Created)
DT 01-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Fibronectin binding protein B precursor.
GN Name=fnbB;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman D2C;
RA Grundmeier M., Hussain M., Becker P., Sinha B.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF629122; CAF32748.1; -.
KW Signal.
FT SIGNAL 1 35 Potential.
FT CHAIN 35 677 fibronectin binding protein B.
SQ SEQUENCE 677 AA; 74737 MW; 37CAD12EFDPC160B1 CRC64;

Query Match 19.1%; Score 515; DB 2; Length 677;
Best Local Similarity 26.8%; Pred. No. 9e-16;
Matches 150; Conservative 95; Mismatches 181; Indels 134; Gaps 18;

QY 1 VGTLLIGFLSSKEADASE---NSVTQSDSASNESKSSDSSVSAAPKTDITNVSDTKTS 57
Db 21 LGTMIVGMGQKEAAASEQNNTTVESGSSATESKA-----SETQTT 63
QY 58 SNTNNGETSAQNPAQOQETQSSSTNATTEETPVTEATTTTNOANTPATTQSSNTNAE 117
Db 64 TNNVNTIDETQSYATSSTEQPSQSTQVTTTEAPKTVQAPKVESTRVLP-----112
QY 118 ELVNQTSNETTNDNTVSSVNSPQNSNAENVSTTQDTSSTE---ATPSNESAPOSTDA 174
Db 113 -----SEKVAOKETITGTVQVDTAQPNS-----133
QY 175 SNKDVVNQAVNTSAPRMR--RAFSLAAVA-----ADAPAAAGTDITNLTNTVTGIDSGT---225
Db 134 -----VSEIKPRMKESTDVTAAVEKEVVEETKATGCTVTK-----VEVEGSEIV 179
```

```
QY 226 -----TVYPHOAGYVYKLVNGESVPNSAVKGDFTKLTVPKELNLNGVSTAKVPPIMAG 278
Db 180 GHKQDTNVVPHNAERVTLKYKKGFGGIGKAGDYDFDTLSDNVETHGISTLRKVPKIKST 239
QY 279 D-QVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGST 337
Db 240 DGQVMATGEIIGERKVRVYTFKVEYQVKKDLTAEISLNLFDIPTVTQKGNQVVEVKLGST 299
QY 338 TANKTVLVY-----EKYGFYNLSIKGTIDQIDKTNNTVROTIIYVNPSSGDNVIAPLTG 392
Db 300 TVSKIFNIQYLGVRDNGW-----VTANGRIDTLNKVDGKFSHFAYMKPNQSSSVTVTG 355
QY 393 NL-----KPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVNPEN---FEDVTSNVITFP 445
Db 356 QVTKGNKPGVN-----NPTVKVYKHIGSDDLAESVYAKLDDVSKFEDVTNMSLDFD 407
QY 446 NPQYKVEFNTPDQITPTPIVYVNGHIDPNSKGDALRSTLYGYSNIIWRSMWDNEV 505
Db 408 TNGYSLNFNNLDQ--SKNYVIKGYGYDSNA-SNLEFQTHLFGYVNYYSNLTWKNV 464
QY 506 AFNNGSGSGDIDK---PVV 522
Db 465 AFYSNNAQGDGDKLKEPII 484

RESULT 12
Q53682 PRELIMINARY; PRT; 940 AA.
AC Q53682;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibronectin binding protein B.
GN Names:fnbB;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=8325-4;
RX MEDLINE=92111475; PubMed=1837266;
RA Jonsson K., Signas C., Muller H.P., Lindberg M.;
RT "Two different genes encode fibronectin binding proteins in
RT Staphylococcus aureus. The complete nucleotide sequence and
RT characterization of the second gene.";
RL Eur. J. Biochem. 202:1041-1048(1991).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; X62992; CAA44726.1; -.
DR PIR; S19702; S19702.
DR HSSP; Q53653; 1N67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0009275; C:cell wall (sensu Gram-positive Bacteria); IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR004237; Fn_bind.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF02986; Fn_bind; 2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK signal; 1.
DR TIGRfams; TIGR01167; LPXTG anchor; 1.
DR TIGRfams; TIGR01168; YsIRK signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 940 AA; 10355 MW; E35FBBCA907AE345 CRC64;

Query Match 19.1%; Score 515; DB 2; Length 940;
Best Local Similarity 26.84; Pred. No. 1.3e-15;
Matches 150; Conservative 95; Mismatches 181; Indels 134; Gaps 18;
```

```
QY 1 VGTILGIFLLSSKEADASE---NSVTQSDASNESKSDSSVSAAKPTDDTNVSDTKTS 57
Db 21 LGTMIIVGMQEKRAAASEQNNTTVEESGSATESKA-----SETQTT 63
QY 58 SNTNNGETSAVQNPAAQOETTQSSSTNATTEETPVTGATTTTNNQANTPATQSSNTNAE 117
Db 64 TNNVNTIDETQSYATSATSEQPSQSTQVTTTEAPKTVQAPKVESRDLP----- 112
QY 118 ELVNTQISNETTFNDTNTVSSVNSPQSNNAENVSTTQDTSTE---ATPSNNEAQAQSDTA 174
Db 113 -----SEKVADEKTTGTQVDIAQPSN----- 133
QY 175 SNKDVVNQAVNTSAPRM-RAFSLAAVA-----ADAPAAGTDITNLTNTVTGIDSGT--- 225
Db 134 -----VSEIKPRMKRSTDTVAEKEVVEETKATGTDVINK-----VEVEEGSEIV 179
QY 226 -----TVYPHOAGYVYKLVNGESVPNSAVKGDFTKLTVPKELNLNGVSTAKVPPIMAG 278
Db 180 GHKQDTNVVPHNAERVTLKYKKGFGGIGKAGDYDFDTLSDNVETHGISTLRKVPKIKST 239
QY 279 D-QVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGST 337
Db 240 DGQVMATGEIIGERKVRVYTFKVEYQVKKDLTAEISLNLFDIPTVTQKGNQVVEVKLGST 299
QY 338 TANKTVLVY-----EKYGFYNLSIKGTIDQIDKTNNTVROTIIYVNPSSGDNVIAPLTG 392
Db 300 TVSKIFNIQYLGVRDNGW-----VTANGRIDTLNKVDGKFSHFAYMKPNQSSSVTVTG 355
QY 393 NL-----KPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVNPEN---FEDVTSNVITFP 445
Db 356 QVTKGNKPGVN-----NPTVKVYKHIGSDDLAESVYAKLDDVSKFEDVTNMSLDFD 407
QY 446 NPQYKVEFNTPDQITPTPIVYVNGHIDPNSKGDALRSTLYGYSNIIWRSMWDNEV 505
Db 408 TNGYSLNFNNLDQ--SKNYVIKGYGYDSNA-SNLEFQTHLFGYVNYYSNLTWKNV 464
QY 506 AFNNGSGSGDIDK---PVV 522
Db 465 AFYSNNAQGDGDKLKEPII 484

RESULT 12
O86489 PRELIMINARY; PRT; 1166 AA.
AC O86489;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sdr E protein.
DE Name=sdr E;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Newman;
RX MEDLINE=99098700; PubMed=9884231;
RX Josefsen E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RX Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AJ005647; CAA06652.1; -.
DR PIR; T28680; T28680.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
```


Q6G6H4	PRELIMINARY;	PRT;	957 AA.
ID	Q6G6H4		
AC	Q6G6H4;		
DT	05-JUL-2004 (TREMBLrel. 27, Created)		
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)		
DE	05-JUL-2004 (TREMBLrel. 27, Last annotation update)		
DE	Fibronectin-binding protein.		
GN	Name=fnbB; ORFNames=SA2387;		
OS	Staphylococcus aureus subsp. aureus MSSA476.		
OS	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
NCBI_TaxID=282459;			
FN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MSA476;		
RA	Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,		
RA	Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,		
RA	Rason N., Bentley S.D., Chillingworth C., Chillingworth T.,		
RA	Churcher C., Clark I., Cotton C., Cronin A., Doggett J., Dowd L.,		
RA	Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,		
RA	James K.D., Lennard N., Line A., Mayes R., Moutle S., Mungall K.,		
RA	Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,		
RA	Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,		
RA	Spratt B.G., Parkhill J.;		
RT	"Complete genomes of two clinical Staphylococcus aureus strains:		
RT	evidence for the rapid evolution of virulence and drug resistance.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).		
CC	-1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by		
CC	an amide bond (by similarity).		
DR	EMBL; BX571857; CAG44201.1; -		
DR	InterPro; IPR004237; Fb bind.		
DR	InterPro; IPR005877; GpcS YSIRK.		
DR	InterPro; IPR001899; Gram_pos_anchor.		
DR	Pfam; PF02986; Fb_bind; 1.		
DR	Pfam; PF00748; Gram_pos_anchor; 1.		
DR	Pfam; PF04650; YSIRK_signal; 1.		
DR	TIGRFAMS; TIGR01167; LPXTG_anchor; 1.		
DR	TIGRFAMS; TIGR01168; YSIRK_signal; 1.		
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.		
KW	Cell wall; Peptidoglycan-anchor.		
SQ	SEQUENCE 957 AA; 105980 MW; 28F8BA7FFDFD3EAF CRC64;		
	Query Match		
	Best Local Similarity 26.5%; Score 494; DB 2; Length 957;		
	Matches 149; Conservative		
QY	1 VGTLLIGGLSSKEADASE--NSVTQSDASNESKNDSSVSAAPKTDITNVSDTKTS 57		
Db	21 LGTIVVMGQEKAAAEQNNITVEEGSGSATERKA-----SETQTT 63		
QY	58 SNTNNGETSVQAQPAQOETTQSSSTNATTETPTVGEATTTTNOANTPATQSSNTNAE 117		
Db	64 TNNVNTIDEQSYSATSTEQPSQSTQVTTTEAPTVAQPKVETSEVDLP----- 112		
QY	118 ELVNQTSNETTFDNTVTVSSVNSQNSTAPNVSNTTQDTSTE---ATPSNNSAPQSTDA 174		
Db	113 -----SEKVADKETGTQVDTAQPSN----- 133		
QY	175 SNKDVMVQAQNTSAPMR-----AFSLAAVAADAPAAAGTDITNQLTNVTVGIDSGT--- 225		
Db	134 -----VSEIKPRMKRSTDVAVTEKEVEEAKATGTDVTSK-----VEVEEGSEIV 179		
QY	226 -----TVYPHOAGVVKLVNGFSPVNSAVKGTFFKITVPKELNLNGVSTAKVPPIMA 277		
Db	180 GHNNKETNVVNPNAERVTLKYWKFDGKIPGDIYFOFTLSNNVETHGISPLRKVPDIKS 239		
QY	278 -GQVILANGVIDSGNVYITFTDYYVNTKDDVKAFLTMTPAYVIDENPVKKTGNVTLATGIGS 336		
Db	240 KDNILAVGKVMDEKIRYFTFDYINNKNLMAELNLNLFIDPTTVTKQGQKQTVVKLGE 299		
QY	337 TTAKTVLVLDY-----EKYKFFNLSIKGTDIQDIDKTNNTYRQTIYNPNPSGDNVIAPULT 391		
Db	300 NKTISKFDIKYLDGVKDNWG-----VTVNGRITLDDKNSKIHHLAYINPKKSDMTSITIN 355		

392 GNLKPTDSNALIDQONTSIKIVKYVDNAD;SSEYFVN---PENFEDVTSVN--ITFFN 444
356 GFFA---KGLYTGNTVTKVYVYLRSDLPESVYANTNDQKFDVNDMSDKLTGE 411
447 PNQYKVENTPDDQITTPYIVVVGCHIDPNSKGLDLALRSTLYGYSNIWR-----SM5W 501
412 NGSYKL---TUDALNKKGYVVSFEKYNENDK-ELLFRNLHGHYHANYGYYYYYPVSLTW 467
502 DNEVAFNNGSGDGDIDKPVVP 523
468 DNGVAFYSNNAQGDGDKNDP 489
RESULT 15
Q99W46 PRELIMINARY; PRT; 1141 AA.
AC Q99W46;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
protein.
GN Name=sdrE; OrderedLocusNames=SA0521;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=158879;
RX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=N315;
RC MEDLINE=21311952; PubMed=11418146;
RX Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Ogasawara N., Hayashi H., Hiramatsu K.,
RA Hattori M., Yagasaki N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancelot 357:1225-1240 (2001).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond [by similarity].
DR ENBL; AP003131; BAB41752.1; -.
DR PIR; E89824; E89824.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_YSRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSRK_signal; 1.
DR TIGRams; TIGR01167; LPXTG_anchor; 1.
DR TIGRams; TIGR01168; YSRK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchored.
SQ SEQUENCE 1141 AA; 124026 MW; 445419D0B8C5A4F8 CRC64;
Query Match 18.3%; Score 494; DB 2; Length 1141;
Best Local Similarity 28.0%; Pred. No. 1.5e-14;
Matches 169; Conservative 92; Mismatches 218; Indels 124; Gaps 26;
QY 1 VGTLLFGLLSKREADASENSVTQSDASNESKNSDSSVSAAPKTDITNVSDPKTSNT 60
DB 37 VGTLLFG-LGNQKAAENSTSE-NAKQDATTSDNKEV-----VSETENNST 84
QY 61 NNGETSAQNPAQOCTTQSSSTNATTEETPVTGEATTTTNTQANTPATTOSSNTNAB-EL 119
DB 85 ENDST---NPKEKNTSDQPEAKEST-----TSSTQOQNVTATTETKPNIEKEN 135
QY 120 VNQTSNETTFNDNTV-----SSVNSPQNSTNAENV--T 152

Db 136 VKPSTDKTATDEDTSVILLEBKAPVNTNDVTTKPBSTSEIQTKPTTTPQBSTNIENSQPOPT 195
QY 153 TQDTSTEATPSNNESAP-----QSTDAENKQVNVQAVNTSAPR 190
Db 196 PSKVDNQVTDATNPKEPVNVSKEELKNNPEKLKELVRNDNNTDRSTKEVATAPTSAPEKR 255
QY 191 MRAPSLAAVAADAPAAAGTDITNQLT-----NVTVGIDSGTIVYPHQAGVYKLNNGYGSFVENS 246
Db 256 LNAKMRFAVAQFAAVASNNVNDLIITVTKQTKVGGKONVAAAHGKDIEYDTEFTDNK 315
QY 247 AVKGTETKI-----TVPKELMLNGVTSTAKVPPIMAGD---QVLANGVID--SDGNVIYTF 297
Db 316 VKKGTMTINVDKNVIPSDL-----TDKNDPIDITDPSGRVIAKGTDFDKATKQIITYTF 368
QY 298 TDYVNTKDDVKATILMPAYIDPENY--KKTGNVTLATGIGSTTANKTVLVDYEKYCKFYN 355
Db 369 TDYVDBKYEIDIKARLTLYSYIDKQAVENETSLMLTATAGKETSQN--VSVDYQDPMVHGD 426
QY 356 LSIKGTIDQIDKNTNTYBQTIYVNP---SGDNVIAPVL-----TGNLKPNTDSNALIDQ 406
Db 427 SNIQSIPTKLDENKQTIQOIYVNPPLKKTATNTKVDIAGSQVDDYGNKLGNGS-TIID- 484
QY 407 QNTSIKVYKVDNAADLSBSYFV-NPENFEDVTSVNIITFPNPNQYKVEFNTPD-DOI TTP 464
Db 485 QNTEIKVYKVNPNQQLPQSNRIYDFSQYEDVTSQ---FDNKKSFNNVATLDFGDI NS 540
QY 465 YIVVWNGHIDPNSKGD-----ALRST-LYGNNSNIIWRSMWDNEVAFNNGSGGDGI 517
Db 541 YIIKVVSXYTPTSDGELDIAQGTSMRTTDKYG-----YNYAGYGNFIVTSNDTGGDGT 595
QY 518 DKP 520
Db 596 VRP 598

Search completed: November 3, 2004, 19:10:17
Job time : 111.231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:48:10 ; Search time 17.8038 Seconds
(without alignments)
1894.005 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_425

Perfect score: 487

Sequence: 1 TLTPAYIDPENVKGTGNT.....NVIAPVLGNLXKNTDSNAL 94

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	DB ID	Description
1	487	100.0	331	5	Aae29263 Staphyloc
2	487	100.0	345	2	Aaw1555 Fibronect
3	487	100.0	496	5	Aau75490 S. aureus
4	487	100.0	496	6	Ada89664 Staphyloc
5	487	100.0	496	6	Ada89663 Staphyloc
6	487	100.0	520	5	Aae29262 Staphyloc
7	487	100.0	927	6	Abm72221 Staphyloc
8	487	100.0	933	3	Aay58435 Staphyloc
9	487	100.0	933	4	Aab69508 Staphyloc
10	487	100.0	933	6	Abj18947 Pathogen
11	487	100.0	936	2	Aaw89801 Staphyloc
12	422	86.7	935	6	Abu16402 Protein e
13	422	86.7	1021	4	Aau33975 Staphyloc
14	422	86.7	1021	4	Aau36951 Staphyloc
15	422	86.7	1018	1	Aap82115 Fibronect
16	422	86.7	1018	4	Aau37245 Staphyloc
17	422	86.7	1018	4	Aau34301 Staphyloc
18	422	86.7	1018	6	Abj18922 Pathogen
19	422	86.7	1018	6	Abm72537 Staphyloc
20	422	86.7	1018	6	Abm79017 Staphyloc
21	422	86.7	1018	6	Abm79016 Staphyloc
22	422	86.7	1018	6	Abm79015 Staphyloc
23	422	86.7	1018	6	Abu42557 Protein e
24	422	86.7	1018	6	Abu42557 Protein e
25	422	86.7	1018	6	ADO84849 S epiderm

ALIGNMENTS

RESULT 1

AAE29263 AAE29263 standard; protein; 331 AA.

XX AC AAE29263;

XX AC (first entry)

DT 27-JAN-2003 (first entry)

XX Staphylococcus aureus Clf33 protein.

XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein;

XX immunological; staphylococcal infection; impetigo; pneumonia; furuncle;

XX septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.

XX Staphylococcus aureus.

XX WO200272600-A2.

XX 19-SEP-2002.

XX 28-JAN-2002; 2002WO-US00296.

XX 26-JAN-2001; 2001US-0264072P.

XX 12-MAR-2001; 2001US-0274611P.

XX 18-JUN-2001; 2001US-0298413P.

XX 30-JUL-2001; 2001US-0308116P.

XX (INHI-) INHIBITEX INC.

XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;

XX WPI; 2002-759834/B2.

XX N-PSDB; AAB46862.

XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for

XX treating or preventing Staphylococcus aureus infection e.g. wound

XX infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in

XX a human or animal.

XX Claim 9; Page 70-72; 80pp; English.

XX The invention relates to monoclonal antibody which binds the clumping

XX factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA

Abp40469 Staphyloc
Aay83171 Cell wall
Aay70120 Staphyloc
Aaw89804 Staphyloc
Aau34403 Staphyloc
Aau37547 Staphyloc
Aar07070 Fibronect
Abj19111 Pathogen
Ada89470 Staphyloc
Abj18923 Pathogen
Abm72536 Staphyloc
Abm79020 Staphyloc
Aaw41602 Staphyloc
Abm79019 Staphyloc
Aau33960 Staphyloc
Aau37093 Staphyloc
Aay08643 S. aureus
Abj18982 Pathogen
Abu43489 Protein e
Abu15854 Protein e

CC animal. These staphylococcal infections include wound infections, sepsis,
 CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
 CC present sequence is Staphylococcus aureus Clf33 protein

XX SQ Sequence 331 AA;

Query Match 100.0%; Score 487; DB 5; Length 331;
 Best Local Similarity 100.0%; Pred. No. 4.6e-46;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTPPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKGFYNLSIKGTIDQIDKTN 60
 Db 113 TLTPPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKGFYNLSIKGTIDQIDKTN 172

Qy 61 NTYRQTIYVNPFGDNVIAPIVLTGNLKPNTDSNAL 94

Db 173 NTYRQTIYVNPFGDNVIAPIVLTGNLKPNTDSNAL 206

RESULT 2

AAW31555
 ID AAW31555 standard; protein; 345 AA.

XX AC AAW31555;

XX DT 27-AUG-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 21-MAY-1998 (first entry)

XX DE Fibronectin-binding MSCRAMM derivative pCF33.

XX KW Fibronectin; pCF33; collagen binding protein; sepsis; infection;

XX KW microbial surface component recognising adhesive matrix molecule; MSCRAMM;

XX KW adhesin; vaccine; immunisation; diagnosis; therapy; mastitis.

XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers

FT Peptide 1..12 /note= "vector pQE30-derived peptide"

XX WO9743314-A2.

XX PD 20-NOV-1997.

XX PF 14-MAY-1997; 97WO-US008210.

XX PR 16-MAY-1996; 96US-0017678P.

XX PA (TEXA) UNIV. TEXAS A & M SYSTEM.

XX PA (UABR-) UAB RES FOUND.

XX PI Hoeek M, Patti JM, House-Pompeo K, Sthanam N, Symersky J;

XX WPI; 1998-008801/01.

XX PT Antibody that interacts with collagen binding domain of Staphylococcal
 PT cna gene product - useful to prevent bacterial sepsis in animal infected
 PT with Staphylococcus aureus.

XX PS Disclosure; Page 91; 143pp; English.

XX CC This protein comprises Staphylococcus aureus fibronectin-binding
 CC microbial surface component recognising adhesive matrix molecule (MSCRAMM)
 CC derivative pCF33, plus a vector-derived N-terminal peptide. The invention
 CC relates to claimed nucleic acid sequences (see AAR3/436-38) encoding S.
 CC aureus collagen binding protein (CBP) epitopes M17, M31 and M55 (see
 CC AAW31552-54) that confer protection against S. aureus infection. CBP
 CC protein and antigenic epitopes are contemplated for use in the treatment
 CC of pathological infections, especially to prevent bacterial adhesion to
 CC collagen. The epitopes are also contemplated for use in the preparation
 CC of vaccines and as carrier proteins in vaccine formulations, as well as
 CC in the formulation of compositions for the prevention of S. aureus

CC infection. pCF33 and pQD (see AAW31556) were used to raise anti-MSCRAMM
 CC polyclonal antibodies used in passive immunisation against bovine
 CC mastitis. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-
 CC AUG-2003 to correct OS field.)

XX SQ Sequence 345 AA;

Query Match 100.0%; Score 487; DB 2; Length 345;

Best Local Similarity 100.0%; Pred. No. 4.9e-46;

Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTPPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKGFYNLSIKGTIDQIDKTN 60

Db 125 TLTPPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKGFYNLSIKGTIDQIDKTN 184

Qy 61 NTYRQTIYVNPFGDNVIAPIVLTGNLKPNTDSNAL 94

Db 185 NTYRQTIYVNPFGDNVIAPIVLTGNLKPNTDSNAL 218

RESULT 3

AAU75490
 ID AAU75490 standard; protein; 496 AA.

XX AC AAU75490;

XX DT 23-APR-2002 (first entry)

XX DE S. aureus antigenic protein associated protein #10.

XX KW Antigenic protein; vaccine; SREX; antibacterial; antiinflammatory;

XX KW dermatological; antiulcer; tuberculostatic; immunosuppressive;

XX KW septicemia; food poisoning; skin disorders; peritonitis; endocarditis;

XX KW tuberculous; blood infection; sepsis; meningitis; pneumonia;

XX KW stomach ulcer; gonorrhea; necrotising fasciitis; impetigo;

XX KW Lyme's disease; gastro-enteritis; dysentery; shigellosis.

XX OS Staphylococcus aureus.

XX FN WO200198499-A1.

XX PD 27-DEC-2001.

XX PF 20-JUN-2001; 2001WO-GB002685.

XX PR 20-JUN-2000; 2000GB-00014907.

XX PA (UYSH-) UNIV SHEFFIELD.

XX PA (BIOS-) BIOSYNEXUS INC.

XX PI Foster S, McDowell P, Brummell K, Clarke S;

XX WPI; 2002-106544/14.

XX PT Identifying antigenic polypeptides expressed by pathogenic organisms
 PT e.g., Staphylococcus aureus during infection, by SREX (serological
 PT identification of antigens by recombinant expression cloning) techniques.

XX PS Disclosure; Page 72-74; 85pp; English.

XX CC The invention relates to a method for identifying antigenic polypeptides
 CC expressed by pathogenic organisms e.g., Staphylococcus aureus during
 CC infection, by SREX (serological identification of antigens by
 CC recombinant expression cloning) techniques. The method involves providing
 CC a nucleic acid library encoding genes/partial gene sequences of
 CC pathogenic organisms, transforming/transfecting the library into host
 CC cells, contacting the polypeptides expressed by the genes with autologous
 CC antisera derived from an animal infected with, or has been infected with
 CC the pathogen and purifying the nucleic acid encoding the polypeptide or
 CC partial polypeptide binding to the antisera. Also included are the
 CC nucleic acids and polypeptides isolated by the method, vectors and
 CC transformed cells expressing them, a vaccine comprising the polypeptide
 CC and the production of monoclonal antibodies against the polypeptides. The

CC protein and vaccine are useful for immunising an animal (preferably
 CC human) against a pathogenic microbe. The proteins and antibodies are
 CC useful for manufacturing a medicament for treating Staphylococcus aureus-
 CC associated septicaemia, food poisoning or skin disorders or
 CC Staphylococcus epidermidis-associated septicaemia, peritonitis,
 CC endocarditis, tuberculosis, blood infections, sepsis, meningitis,
 CC pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, impetigo,
 CC Lyme's disease, gastro-enteritis, dysentery and shigellosis. The present
 CC sequence is an S. aureus protein sequence associated with the antigenic
 CC proteins of the invention. Note: The present sequence is included in the
 CC sequence listing but is not mentioned anywhere else in the specification
 CC
 CC Sequence 496 AA;

Query Match 100.0%; Score 487; DB 5; Length 496;
 Best Local Similarity 100.0%; Pred. No. 7.8e-46;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKGYKGYFYNLSIKGTIDQIDKTN 60
 Db 332 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKGYKGYFYNLSIKGTIDQIDKTN 391
 QY 61 NTVRQIYVNPNGDNVIAPLVTGNLKPNTDSNAL 94
 Db 392 NTVRQIYVNPNGDNVIAPLVTGNLKPNTDSNAL 425

RESULT 4
 ADA89664
 ID ADA89664 standard; protein; 496 AA.
 AC
 AC ADA89664;
 XX

20-NOV-2003 (first entry)

Staphylococcus aureus antigenic protein #203.

CC antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
 CC antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
 CC antiulcer; immunostimulant; ophthalmological; pathogenic microbe;
 CC bacteraemia; septic shock; organ infection; skin infection;
 CC bacterial basal colonisation; bacterial eye infection; septicaemia;
 CC tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
 CC sepsis; meningitis; pneumonia; stomach ulcer; gonorrhea; toxic shock;
 CC necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
 CC gastro-enteritis; dysentery; shigellosis; skin disorder.

Staphylococcus aureus

WO2003011899-A2.

13-FEB-2003.

02-AUG-2002; 2002WO-GB003606.

02-AUG-2001; 2001GB-00018825.

09-JAN-2002; 2002GB-00000349.

(UYSH-) UNIV SHEFFIELD.

(BIOS-) BIOSYNEXUS INC.

Foster S, Mond J, Clarke S, McDowell P, Brummel K;

WPI; 2003-256434/25.

CC New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
 CC useful as a vaccine for immunizing humans against e.g. bacteremia, septic
 CC shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhea or
 CC impetigo.

Claim 4; Page 161; 189pp; English.

The present invention describes an antigenic protein or its part, which

CC is for use as a vaccine. The antigenic protein is encoded by an isolated
 CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
 CC sequences (designated dnaSA and dna SE, respectively; and which encodes a
 CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
 CC composition comprising at least one antigenic protein; (2) a method of
 CC immunising an animal against a disease or condition caused by a
 CC pathogenic microbe by administering the antigenic protein or the vaccine;
 CC (3) an antibody or its binding part obtainable by the method above; (4)
 CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
 CC hybridoma cell line produced by the method of (4); and (6) identifying
 CC opsonic antigens expressed by a pathogenic microbe. The antigenic
 CC proteins have antibacterial, neuroprotective, immunosuppressive,
 CC antiinflammatory, antiulcer, immunostimulant and ophthalmological
 CC activities, and can be used in vaccines. The antigenic proteins or
 CC vaccines can be used for immunising an animal (specifically a human)
 CC against a disease or condition caused by a pathogenic microbe, e.g.
 CC bacteraemia, septic shock, organ infection, skin infection, bacterial
 CC basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
 CC bacteria-associated food poisoning, blood infections, peritonitis,
 CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhea,
 CC strep throat, streptococcal-associated toxic shock, necrotising
 CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
 CC dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
 CC skin disorders, S. epidermidis-associated septicaemia, peritonitis or
 CC endocarditis. The present sequence represents a S. aureus antigenic
 CC protein sequence from the present invention.

CC Sequence 496 AA;

Query Match 100.0%; Score 487; DB 6; Length 496;
 Best Local Similarity 100.0%; Pred. No. 7.8e-46;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKGYKGYFYNLSIKGTIDQIDKTN 60
 Db 332 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKGYKGYFYNLSIKGTIDQIDKTN 391
 QY 61 NTVRQIYVNPNGDNVIAPLVTGNLKPNTDSNAL 94
 Db 392 NTVRQIYVNPNGDNVIAPLVTGNLKPNTDSNAL 425

RESULT 5

ADA89663

ID ADA89663 standard; protein; 496 AA.

AC ADA89663;

20-NOV-2003 (first entry)

Staphylococcus aureus antigenic protein #202.

CC antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
 CC antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
 CC antiulcer; immunostimulant; ophthalmological; pathogenic microbe;
 CC bacteraemia; septic shock; organ infection; skin infection;
 CC bacterial basal colonisation; bacterial eye infection; septicaemia;
 CC tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
 CC sepsis; meningitis; pneumonia; stomach ulcer; gonorrhea; toxic shock;
 CC necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
 CC gastro-enteritis; dysentery; shigellosis; skin disorder.

Staphylococcus aureus.

WO2003011899-A2.

13-FEB-2003.

02-AUG-2002; 2002WO-GB003606.

02-AUG-2001; 2001GB-00018825.

09-JAN-2002; 2002GB-00000349.

PA (UYSH-) UNIV SHEFFIELD.
PA (BIOS-) BIOSYNEXUS INC.
XX
PI Foster S, Mond J, Clarke S, McDowell P, Brummel K;
XX WPI; 2003-256434/25.
DR
XX
XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
XX useful as a vaccine for immunizing humans against e.g. bacteremia, septic
XX shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhea or
XX impetigo.
XX
XX Claim 4; Page 160-161; 189pp; English.
XX
XX The present invention describes an antigenic protein or its part, which
XX is for use as a vaccine. The antigenic protein is encoded by an isolated
XX DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
XX sequences (designated dnaSA and dna SE, respectively; and which encodes a
XX protein expressed by a pathogenic organism. Also described: (1) a vaccine
XX composition comprising at least one antigenic protein; (2) a method of
XX immunising an animal against a disease or condition caused by a
XX pathogenic microbe by administering the antigenic protein or the vaccine;
XX (3) an antibody or its binding part obtainable by the method above; (4)
XX preparing a hybridoma cell line producing monoclonal antibodies; (5) a
XX hybridoma cell line produced by the method of (4); and (6) identifying
XX opsonic antigens expressed by a pathogenic microbe. The antigenic
XX proteins have antibacterial, neuroprotective, immunosuppressive,
XX antiinflammatory, antiulcer, immunostimulant and ophthalmological
XX activities, and can be used in vaccines. The antigenic proteins or
XX vaccines can be used for immunising an animal (specifically a human)
XX against a disease or condition caused by a pathogenic microbe, e.g.
XX bacteraemia, septic shock, organ infection, skin infection, bacterial
XX basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
XX bacteria-associated food poisoning, blood infections, peritonitis,
XX endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
XX strep throat, streptococcal-associated toxic shock, necrotising
XX fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
XX dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
XX skin disorders, S. epidermidis-associated septicaemia, peritonitis or
XX endocarditis. The present sequence represents a S. aureus antigenic
XX protein sequence from the present invention.
SQ Sequence 496 AA;
Query Match 100.0%; Score 487; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 7.8e-46;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVYDYKYGKFNLSIKGTIDQIDKTN 60
Db 332 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVYDYKYGKFNLSIKGTIDQIDKTN 391
QY 61 NTYQTIVYVPSGDNVIAPIVLTGNLKPNTDSNAL 94
Db 392 NTYQTIVYVPSGDNVIAPIVLTGNLKPNTDSNAL 425
RESULT 6
AAE29262
ID AAE29262 standard; protein; 520 AA.
XX
XX AAE29262;
XX
XX 27-JAN-2003 (first entry)
XX
XX Staphylococcus aureus Clf40 protein.
XX
XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
XX immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
XX septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.
XX
XX Staphylococcus aureus.
XX
XX
PN WO200272600-A2.
XX
XX 19-SEP-2002.
XX
XX 28-JAN-2002; 2002WO-US002296.
XX
XX 26-JAN-2001; 2001US-0264072P.
XX 12-MAR-2001; 2001US-0274611P.
XX 18-JUN-2001; 2001US-0298413P.
XX 30-JUL-2001; 2001US-0308116P.
XX
XX (INHI-) INHIBITEX INC.
XX
XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
XX WPI; 2002-759834/82.
XX N-FSDB; AAD48661.
XX
XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for
XX treating or preventing Staphylococcus aureus infection e.g. wound
XX infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
XX a human or animal.
XX
XX Claim 9; Page 67-69; 80pp; English.
XX
XX The invention relates to monoclonal antibody which binds the clumping
XX factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
XX monoclonal antibody is useful for treating or preventing S. aureus
XX infection in a human or animal, and for inhibiting the binding of
XX staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
XX S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3
XX protein is useful for inducing an immunological response in a human or
XX animal. These staphylococcal infections include wound infections, sepsis,
XX impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
XX present sequence is Staphylococcus aureus Clf40 protein
XX
SQ Sequence 520 AA;
Query Match 100.0%; Score 487; DB 5; Length 520;
Best Local Similarity 100.0%; Pred. No. 8.3e-46;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVYDYKYGKFNLSIKGTIDQIDKTN 60
Db 293 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVYDYKYGKFNLSIKGTIDQIDKTN 352
QY 61 NTYQTIVYVPSGDNVIAPIVLTGNLKPNTDSNAL 94
Db 353 NTYQTIVYVPSGDNVIAPIVLTGNLKPNTDSNAL 386
RESULT 7
ABM72221
ID ABM72221 standard; protein; 927 AA.
XX
XX ABM72221;
XX
XX 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus protein #1461.
XX
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX enzymatic assay; antibiotic target.
XX
XX Staphylococcus aureus.
XX
XX WO200294868-A2.
XX
XX 28-NOV-2002.
XX
XX 27-MAR-2002; 2002WO-IB0002637.
XX
XX 27-MAR-2001; 2001GB-00007661.

XX (CHIR-) CHIRON SPA.
XX Masignani V, Mora M, Scarselli M;
XX WPI; 2003-120786/11.
XX N-PSDB; ACF73781.
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.
XX Claim 1; SEQ ID NO 2922; 49pp; English.
XX
XX The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is
XX useful as a pharmaceutical, particularly as a vaccine for treating or
XX preventing infection due to Staphylococcus bacteria, specifically an
XX infection caused by S. aureus. The composition is particularly useful for
XX treating or preventing sepsis in a patient. The composition can also be
XX used for diagnostics. The protein is also used in an assay for enzymatic
XX studies and as a target for antibiotics. This sequence represents one of
XX the novel S. aureus proteins of the invention
XX
XX Sequence 927 AA;
Query Match 100.0%; Score 487; DB 6; Length 927;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 391
QY 61 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 94
DB 392 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 425
RESULT 8
ID AAY58435 standard; protein; 933 AA.
XX AAY58435;
XX 27-MAR-2000 (first entry)
XX Staphylococcus aureus fibrinogen binding Clfa protein.
XX Clfa; fibrinogen binding protein; bacterial colonisation;
XX indwelling medical device; staphylococcal infection.
XX Staphylococcus aureus.
XX Key Location/Qualifiers
XX Peptide 1..39
XX /note= "Signal peptide"
XX Region 40..559
XX /note= "Region A"
XX Region 332..550
XX /note= "Fibrinogen-binding region"
XX Region 560..867
XX /note= "Region R"
XX Region 896..900
XX /note= "Gram positive wall-associated consensus motif"
XX US6008341-A.
XX 28-DEC-1999.
XX 22-AUG-1994; 94US-00293728.
XX

PR 22-AUG-1994; 94US-00293728.
XX (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX Foster TJ, Mcdevitt DL;
XX WPI; 2000-096389/08.
XX N-PSDB; AA255832.
XX Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus
XX aureus, useful for treatment or prevention of infections.
XX Disclosure; Fig 2A-1-4; 35pp; English.
XX
XX This sequence represents the Staphylococcus aureus fibrinogen-binding
XX protein, Clfa. Clfa is an important receptor involved in S. aureus
XX colonisation of indwelling medical devices (e.g., catheters, artificial
XX heart valves). Shortly after implantation, the surfaces of medical
XX devices become coated with host plasma and matrix proteins such as
XX fibrinogen and fibronectin, and there is considerable evidence to suggest
XX that bacterial adherence to fibrinogen/fibrin is important in the
XX initiation of device-related infection. The fibrinogen-binding region of
XX Clfa is thought to reside between residues 332 and 550 in a region
XX designated A. The protein also contains a repeated region (region R)
XX comprising 154 repeats of the dipeptide Ser-Asp, and the C-terminus
XX contains features present in surface proteins of other Gram positive
XX bacteria that are responsible for anchoring the protein to the cell wall
XX and cell membrane. Clfa, or its fragments, may be used to block S. aureus
XX colonisation of wounds, to prevent adherence of S. aureus to indwelling
XX medical devices, as vaccines to protect against S. aureus infection
XX (e.g., mastitis in ruminants), to raise specific antibodies, and for
XX diagnosis (by agglutination or immunoassay). The specific antibodies are
XX used for passive immunisation, to block infection of wounds or adhesion
XX of S. aureus and for diagnosis. Nucleotides encoding Clfa and its
XX fragments may be used as diagnostic probes
XX
XX Sequence 933 AA;
Query Match 100.0%; Score 487; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 391
QY 61 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 94
DB 392 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 425
RESULT 9
ID AAB69508 standard; protein; 933 AA.
XX AAB69508;
XX 23-APR-2001 (first entry)
XX Staphylococcus aureus Clfa protein.
XX Staphylococcus aureus; clfa; antibiotic; vaccine;
XX fibrinogen binding protein; bacterial infection; mastitis.
XX Staphylococcus aureus.
XX US6177084-B1.
XX 23-JAN-2001.
XX 19-OCT-1999; 99US-00421868.
XX 22-AUG-1994; 94US-00293728.
XX

XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX Foster TJ, Mcdevitt DL;
XX WPI; 2001-181608/18.
XX DR N-PSDB; AAF58593.
XX
XX Novel Staphylococcus aureus fibrinogen binding protein useful as vaccine
XX to prevent infection, promote wound healing, block adherence to
XX indwelling medical devices and for diagnosing staphylococcus aureus
XX infection.
XX
XX Claim 5; Fig 2; 30pp; English.
XX
XX The present sequence is a novel Staphylococcus aureus fibrinogen binding
XX protein. It is useful as a vaccine to protect against human and animal
XX infections caused by S.aureus, such as against mastitis, to block
XX S.aureus from colonising and infecting a wound, to block adherence of
XX S.aureus to indwelling medical devices such as catheters, replacement
XX heart valves and cardiac assist devices. The protein can be used to
XX diagnose bacterial infections. Polyclonal and monoclonal antibodies
XX raised against the fibrinogen binding protein are useful for passively
XX immunising against infections caused by S.aureus, to prevent infection of
XX a wound and to diagnose bacterial infections
XX
XX Sequence 933 AA;
XX
Query Match 100.0%; Score 487; DB 4; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVTLATGIGSTANKTVLDVYKGYKGFYNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVTLATGIGSTANKTVLDVYKGYKGFYNLSIKGTIDQIDKTN 391
QY 61 NTRYQTIYVNPFGDNVIAPIVLTGNLKPNTDSNAL 94
DB 392 NTRYQTIYVNPFGDNVIAPIVLTGNLKPNTDSNAL 425
RESULT 10
ID ABJ18947
AC ABJ18947 standard; protein; 933 AA.
XX
XX
XX 06-MAR-2003 (first entry)
XX
XX Pathogen specific antigen related staphylococcal protein SEQ ID No 93.
XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
XX hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
XX auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
XX autoimmune disease; HIV; hepatitis.
XX Staphylococcus sp.
XX WO200259148-A2.
XX
XX 01-AUG-2002.
XX
XX 21-JAN-2002; 2002WO-EP000546.
XX
XX 26-JAN-2001; 2001AT-00000130.
XX
XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX Meinke A, Nagy E, Von Ahlsen U, Klade C, Henicks T, Zauner W;
XX Mihh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;
XX Tempelmaier B;
XX WPI; 2003-075410/07.
XX

XX Identifying, isolating and producing hyperimmune serum-reactive antigens
XX from a pathogen, for preparing vaccine or medicament for treating or
XX preventing e.g. staphylococcal infections, comprises providing antibody
XX preparation.
XX
XX Example 7; Page 160; 252pp; English.
XX
XX The invention relates to a novel method for identifying, isolating and
XX producing hyperimmune serum-reactive antigens from a pathogen, tumour,
XX allergen, a tissue or host prone to auto-immunity, where the antigens are
XX used in a vaccine, comprises providing antibody preparation from a plasma
XX pool of a type of animal, or individual sera with antibodies against the
XX specific pathogen, tumour, allergen, tissue or host prone to auto-
XX immunity. The hyperimmune serum-reactive antigens comprising any of the
XX 62 sequences of 53-2261 amino acids fully defined in the specification,
XX or their hyperimmune fragments are useful for the manufacture of a
XX pharmaceutical preparation, particularly a vaccine against staphylococcal
XX infections or colonisation against S. aureus or S. epidermidis. The
XX preparation of antibodies is useful for the manufacture of a medicament
XX for treating or preventing staphylococcal infections or colonisation
XX against S. aureus or S. epidermidis. The antibody preparations may also
XX be used for diagnostic and imaging purposes. Other conditions that can be
XX treated include cancer, autoimmune diseases or infections caused by viral
XX (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
XX sequence represents a staphylococcal protein relating to the method for
XX identifying and producing pathogen specific antigens of the invention
XX
XX Sequence 933 AA;
XX
Query Match 100.0%; Score 487; DB 6; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVTLATGIGSTANKTVLDVYKGYKGFYNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVTLATGIGSTANKTVLDVYKGYKGFYNLSIKGTIDQIDKTN 391
QY 61 NTRYQTIYVNPFGDNVIAPIVLTGNLKPNTDSNAL 94
DB 392 NTRYQTIYVNPFGDNVIAPIVLTGNLKPNTDSNAL 425
RESULT 11
ID AAW89801
XX AAW89801 standard; protein; 936 AA.
XX
XX AAW89801;
XX
XX 16-MAR-1999 (first entry)
XX Staphylococcus aureus protein SEQ ID #5249.
XX
XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
XX cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
XX skin infection; surgical wound infection; scalded skin syndrome;
XX toxic shock syndrome.
XX Staphylococcus aureus.
XX EP786519-A2.
XX
XX 30-JUL-1997.
XX
XX 07-JAN-1997; 97EP-00100117.
XX
XX 05-JAN-1996; 96US-0009861P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Kunsch CA, Choi GH, Barash SC, Dillion PJ, Fannon MR, Rosen CA;
XX WPI; 1997-374922/35.
XX

XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -
 PT stored on computer readable medium and used in the production of anti-
 PT *S. aureus* vaccines.
 XX
 PS Claim 23; Page 3255-3258; 3271pp; English.
 XX
 CC This sequence represents a *Staphylococcus aureus* protein sequence of the
 CC invention. The DNA sequences encoding the *S. aureus* proteins are recorded
 CC on a computer readable medium, preferably selected from a floppy or hard
 CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.
 CC Homology searches using the *S. aureus* DNA sequences allows putative
 CC functions to be assigned so that protein-encoding or regulatory regions
 CC of commercial, therapeutic or industrial importance can be obtained.
 CC Specifically, sequences which are likely to encode antigens have been
 CC identified and these polypeptides can be used in a vaccine composition
 CC against *S. aureus* infection. The polypeptides can also be used in a kit
 CC for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated
 CC in numerous human diseases, including cellulitis, eyelid infections, food
 CC poisoning, osteomyelitis, skin and surgical wound infections, scalded
 CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the
 CC DNA sequences can be used for recombinant production of the polypeptides.
 CC The new DNA sequences (and their fragments) are useful as primers or
 CC probes for isolating homologues of any of the 5191 *S. aureus* DNA sequences
 CC contained on the computer readable medium
 XX
 XX Sequence 936 AA;

Query Match 100.0%; Score 487; DB 2; Length 936;
 Best Local Similarity 100.0%; Pred. No. 1.8e-45;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTPAYIDPENVKKTGNVTLAGTGTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
 DB 341 TLTPAYIDPENVKKTGNVTLAGTGTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 400
 QY 61 NTRYQTIYVNPNGDNVIAPIVLTGNLKPDTSNAL 94
 DB 401 NTRYQTIYVNPNGDNVIAPIVLTGNLKPDTSNAL 434

RESULT 12
 ABU16402
 ID ABU16402 standard; protein; 935 AA.
 XX
 AC ABU16402;
 XX
 DT 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #1929.
 DE
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW
 XX *Staphylococcus aureus*.
 OS
 XX WO200277183-A2.
 XX
 XX 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US0009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.

DR N-PSDB; ACA20272.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 44326; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 935 AA;

Query Match 86.7%; Score 422; DB 6; Length 935;
 Best Local Similarity 84.9%; Pred. No. 3.6e-38;
 Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
 QY 2 LTMPAYIDPENVKKTGNVTLAGTGTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 61
 DB 333 LTMPAYIDPENVKKTGNVTLAGTGTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 392
 QY 62 TYRQTIYVNPNGDNVIAPIVLTGNLKPDTSNAL 94
 DB 393 TYRQTIYVNPNGDNVIAPIVLTGNLKPDTSNAL 425

RESULT 13
 AAU33975
 ID AAU33975 standard; protein; 1021 AA.
 XX
 AC AAU33975;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE *Staphylococcus aureus* cellular proliferation protein #251.
 XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 XX *Staphylococcus aureus*.
 OS
 XX WO200170955-A2.
 PN
 XX 27-SEP-2001.


```
XX 21-MAR-2001; 2001WO-US009180.
PF 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS51834.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 5471; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1021 AA;
XX
XX Query Match 86.7%; Score 422; DB 4; Length 1021;
XX Best Local Similarity 84.9%; Pred. No. 4e-38;
XX Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
XX
XX QY 2 LTMPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 61
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 333 ITMPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 392
XX
XX QY 62 TYRQTIYVPSGDNVIAPLVLTGNLKPNTDSNAL 94
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 393 TYRQTIYVPSGDNVIAPLVLTGNLKPNTDSNAL 425
XX
XX Db
XX
XX RESULT 14
XX AAU36951
XX ID AAU36951 standard; protein; 1021 AA.
XX AC AAU36951;
XX
XX XX 14-FEB-2002 (first entry)
XX
XX Staphylococcus aureus cellular proliferation protein #1121.
XX
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX
```

```
FN WO200170955-A2.
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
XX 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS54810.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 12544; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1021 AA;
XX
XX Query Match 86.7%; Score 422; DB 4; Length 1021;
XX Best Local Similarity 84.9%; Pred. No. 4e-38;
XX Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
XX
XX QY 2 LTMPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 61
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 333 ITMPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 392
XX
XX Db
XX
XX QY 62 TYRQTIYVPSGDNVIAPLVLTGNLKPNTDSNAL 94
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 393 TYRQTIYVPSGDNVIAPLVLTGNLKPNTDSNAL 425
XX
XX Db
XX
XX RESULT 15
XX AAP82115
XX ID AAP82115 standard; protein; 1018 AA.
XX AC AAP82115;
XX
XX 25-MAR-2003 (revised)
XX 05-JAN-1990 (first entry)
XX
XX Fibronectin binding protein.
XX
XX Fibronectin binding protein; S. aureus; immunisation; mastitis; ruminants;
XX
```

Search completed: November 3, 2004, 19:03:03
Job time : 18.8038 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:54:26 ; Search time 4.60992 Seconds
(without alignments)
1352.277 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_425

Perfect score: 487

Sequence: 1 TLTPAYIDPENVKKTGNVT.....NVIAPVLGNLKPNTDSNAL 94

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	100.0	345	US-08-856-253-7	Sequence 7, Appli
2	487	100.0	933	US-08-293-728-2	Sequence 2, Appli
3	487	100.0	933	US-09-421-868-2	Sequence 2, Appli
4	487	100.0	936	US-08-956-171B-5249	Sequence 5249, Ap
5	487	100.0	936	US-08-781-986A-5249	Sequence 5249, Ap
6	113	23.2	1027	US-08-956-171B-5254	Sequence 5254, Ap
7	113	23.2	1027	US-08-781-986A-5254	Sequence 5254, Ap
8	102.5	21.0	930	US-09-134-001C-5314	Sequence 5314, Ap
9	102.5	21.0	930	US-09-386-962C-10	Sequence 5252, Ap
10	100.5	20.6	251	US-08-956-171B-5252	Sequence 5252, Ap
11	100.5	20.6	251	US-08-781-986A-5252	Sequence 5252, Ap
12	99.5	20.4	582	US-09-147-405B-13	Sequence 13, Appli
13	99.5	20.4	593	US-09-147-405B-11	Sequence 11, Appli
14	99.5	20.4	1092	US-09-147-405B-15	Sequence 15, Appli
15	95	19.5	1132	US-09-348-796A-15026	Sequence 15026, A
16	88.5	18.2	1166	US-09-200-650B-7	Sequence 7, Appli
17	81.5	16.7	1315	US-09-200-650B-5	Sequence 5, Appli
18	79	16.2	1155	US-09-710-279-1780	Sequence 1780, Ap
19	79	16.2	1742	US-09-386-962C-4	Sequence 4, Appli
20	79	16.2	1742	US-09-386-959-4	Sequence 4, Appli
21	73.5	15.1	381	US-09-569-611C-47	Sequence 47, Appli
22	73.5	15.1	400	US-09-569-611C-48	Sequence 48, Appli
23	73.5	15.1	571	US-09-569-611C-48	Sequence 48, Appli
24	73.5	15.1	640	US-09-569-611C-45	Patent No. 5200183
25	73.5	15.1	659	US-09-569-611C-45	Sequence 45, Appli
26	73.5	15.1	668	US-09-569-611C-46	Sequence 46, Appli
27	73.5	15.1	668	US-08-445-050-9	Sequence 9, Appli
28	73.5	15.1	668	US-08-445-050-9	Sequence 9, Appli

28	73.5	15.1	722	1	US-08-347-718B-1	Sequence 1, Appli
29	73.5	15.1	722	1	US-08-445-050-3	Sequence 3, Appli
30	73.5	15.1	722	1	US-08-445-050-7	Sequence 7, Appli
31	73.5	15.1	722	2	US-08-482-262-1	Sequence 1, Appli
32	73.5	15.1	722	2	US-08-204-691-3	Sequence 3, Appli
33	73.5	15.1	722	2	US-08-204-691-7	Sequence 7, Appli
34	73.5	15.1	722	4	US-09-355-295B-4	Sequence 4, Appli
35	73.5	15.1	722	6	5200183-3	Patent No. 5200183
36	73.5	15.1	742	1	US-08-347-718B-2	Sequence 2, Appli
37	73.5	15.1	742	2	US-08-482-262-2	Sequence 2, Appli
38	73.5	15.1	742	6	5200183-2	Patent No. 5200183
39	73.5	15.1	745	1	US-08-445-050-2	Sequence 2, Appli
40	73.5	15.1	745	2	US-08-204-691-2	Sequence 2, Appli
41	73.5	15.1	745	3	US-08-370-223-13	Sequence 13, Appli
42	73.5	15.1	745	4	US-09-355-295B-3	Sequence 3, Appli
43	73.5	15.1	747	3	US-09-347-878-32	Sequence 32, Appli
44	73	15.0	2780	4	US-10-220-587-2	Sequence 2, Appli
45	71	14.6	918	4	US-09-200-650B-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0 Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,253
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,678
FILING DATE: 16-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/SOCKET NUMBER: JANK:193
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-856-253-7

Query Match 100.0%; Score 487; DB 3; Length 345;
Best Local Similarity 100.0%; Pred. No. 1e-51;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
Db 125 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 184

QY 61 NTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNAL 94
Db 185 NTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNAL 218

RESULT 2
US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match 100.0%; Score 487; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 4.1e-51;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
Db 332 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 391

QY 61 NTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNAL 94
Db 392 NTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNAL 425

RESULT 3
US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421,868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match 100.0%; Score 487; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 4.1e-51;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
Db 332 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 391

QY 61 NTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNAL 94

Db 392 NTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNAL 425

RESULT 4

US-08-956-171E-5249
; Sequence 5249, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi

Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Mark S. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5249:
US-08-956-171E-5249

Query Match 100.0%; Score 487; DB 4; Length 936;
Best Local Similarity 100.0%; Pred. No. 4.2e-51;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
Db 341 TLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 400

QY 61 NTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNAL 94
Db 401 NTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNAL 434

RESULT 5
US-08-781-986A-5249
; Sequence 5249, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch


```

; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5254:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1027 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5254

Query Match 23.2%; Score 113; DB 4; Length 1027;
Best Local Similarity 25.0%; Pred. No. 6.6e-05;
Matches 23; Conservative 24; Mismatches 43; Indels 2; Gaps 2;

QY 2 LTMPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDK 60
Db LEINLFDPKTVQINGNQTITSLNEEQTSKELDVKKDGIGNYV-ANLNGSIETFNKAN 368

QY 61 NTYRQTIYVNPNGVNIAPVLVTGNLKPNTDSN 92
Db NRFSHVAFIKENNGKTTSTVTGTLMKGSNQ 400

RESULT 8
US-09-134-001C-5314
; Sequence 5314, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5314
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-5314

Query Match 21.0%; Score 102.5; DB 3; Length 930;
Best Local Similarity 32.9%; Pred. No. 0.0011;
Matches 24; Conservative 13; Mismatches 31; Indels 5; Gaps 2;

QY 2 LTMPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDK 58
Db LKLTSYDKSKVPNNNTKLDVEYKTLSS--VNKTITVEYQKPNENRTANLQSMFTNIDT 443

QY 59 TNNYRQTIYVNP 71
Db 444 KNHTVEQTIYINP 456

RESULT 9
US-09-386-962C-10
; Sequence 10, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; CURRENT FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25

; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-386-962C-10

Query Match 21.0%; Score 102.5; DB 4; Length 930;
Best Local Similarity 32.9%; Pred. No. 0.0011;
Matches 24; Conservative 13; Mismatches 31; Indels 5; Gaps 2;

QY 2 LTMPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDK 58
Db LKLTSYDKSKVPNNNTKLDVEYKTLSS--VNKTITVEYQKPNENRTANLQSMFTNIDT 443

QY 59 TNNYRQTIYVNP 71
Db 444 KNHTVEQTIYINP 456
```

```

; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-386-962C-10

Query Match 21.0%; Score 102.5; DB 4; Length 930;
Best Local Similarity 32.9%; Pred. No. 0.0011;
Matches 24; Conservative 13; Mismatches 31; Indels 5; Gaps 2;

QY 2 LTMPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDK 58
Db LKLTSYDKSKVPNNNTKLDVEYKTLSS--VNKTITVEYQKPNENRTANLQSMFTNIDT 443

QY 59 TNNYRQTIYVNP 71
Db 444 KNHTVEQTIYINP 456

RESULT 10
US-08-956-171E-5252
; Sequence 5252, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5252:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5252:
; US-08-956-171E-5252

Query Match 20.6%; Score 100.5; DB 4; Length 251;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 19:03:11 ; Search time 1.37768 Seconds
(without alignments)
1078.240 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_425

Perfect score: 487

Sequence: 1 TLTPPAYIDPENVKKTGNVT.....NVIAPVLTLGNKPTDSNAL 94

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 66696 seqs, 15802848 residues

Total number of hits satisfying chosen parameters: 66696

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents_AA_New.*
- 1: /cgn2_6/prodata/2/paa/PCT_NEW_COMB.pep.*
 - 2: /cgn2_6/prodata/2/paa/US06_NEW_COMB.pep.*
 - 3: /cgn2_6/prodata/2/paa/US07_NEW_COMB.pep.*
 - 4: /cgn2_6/prodata/2/paa/US08_NEW_COMB.pep.*
 - 5: /cgn2_6/prodata/2/paa/US09_NEW_COMB.pep.*
 - 6: /cgn2_6/prodata/2/paa/US10_NEW_COMB.pep.*
 - 7: /cgn2_6/prodata/2/paa/US11_NEW_COMB.pep.*
 - 8: /cgn2_6/prodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	15.2	94	US-10-220-366A-24254	Sequence 24254, A
2	67.5	13.9	679	US-10-805-684-105	Sequence 105, App
3	64	13.1	1009	US-10-837-269-266	Sequence 266, App
4	64	13.1	1009	US-10-837-269-272	Sequence 272, App
5	64	13.1	1009	US-10-837-269-278	Sequence 278, App
6	64	13.1	1009	US-10-837-269-284	Sequence 284, App
7	64	13.1	1037	US-10-837-269-13	Sequence 13, Appl
8	64	13.1	1037	US-10-837-269-19	Sequence 19, Appl
9	64	13.1	1037	US-10-837-269-25	Sequence 25, Appl
10	64	13.1	1037	US-10-837-269-28	Sequence 28, Appl
11	64	13.1	1037	US-10-837-269-36	Sequence 36, Appl
12	64	13.1	1037	US-10-837-269-39	Sequence 39, Appl
13	64	13.1	1037	US-10-837-269-264	Sequence 264, App
14	64	13.1	1037	US-10-837-269-264	Sequence 264, App
15	64	13.1	1048	US-10-837-269-15	Sequence 15, Appl
16	64	13.1	1048	US-10-837-269-26	Sequence 26, Appl
17	64	13.1	1048	US-10-837-269-37	Sequence 37, Appl
18	64	13.1	1048	US-10-837-269-270	Sequence 270, App
19	64	13.1	1340	US-10-837-269-17	Sequence 17, Appl
20	64	13.1	1340	US-10-837-269-27	Sequence 27, Appl
21	64	13.1	1340	US-10-837-269-38	Sequence 38, Appl
22	64	13.1	1340	US-10-837-269-276	Sequence 276, A
23	64	13.1	2228	US-10-732-923-22722	Sequence 22722, A
24	63.5	13.0	360	US-10-967-702-251	Sequence 251, App
25	62	12.7	1286	US-10-732-923-495	Sequence 495, App

ALIGNMENTS

RESULT 1

US-10-220-366A-24254
; Sequence 24254, Application US/10220366A
; GENERAL INFORMATION:
; APPLICANT: HYSEQ, INC
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-042
; CURRENT APPLICATION NUMBER: US/10/220,366A
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27802
; SOFTWARE: Custom
; SEQ ID NO 24254
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-366A-24254

Query Match 15.2%; Score 74; DB 6; Length 94;
Best Local Similarity 25.7%; Pred. No. 0.13;
Matches 19; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 14 KKTGNVT-LATIGSTTANKTVLVDEYKYGKFNLSIKGTIDDKNTTYRQTYVNFS 72
Db 5 KKKGPNKIRSGKGTDTTNTVTIIRDYGHYTYNKLK-NLEEVDFLDYLPFRNQRE 63

QY 73 GDNVAPVLTLGNLK 86

Db 64 IENMQPITSNEIE 77

RESULT 2

US-10-805-684-105
; Sequence 105, Application US/10805684
; GENERAL INFORMATION:
; APPLICANT: SAKAMOTO, TAKESHI
; APPLICANT: TAKEDA, SHIZU
; TITLE OF INVENTION: PHOS-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: 58748(70342)
; CURRENT APPLICATION NUMBER: US/10/805,684
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/455,766
; PRIOR FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/459,936
; PRIOR FILING DATE: 2003-04-02

Sequence 4727, Ap
Sequence 13054, A
Sequence 18303, A
Sequence 1047, Ap
Sequence 19147, A
Sequence 19146, A
Sequence 1, Appli
Sequence 19212, A
Sequence 265, App
Sequence 271, App
Sequence 277, App
Sequence 283, App
Sequence 253, App
Sequence 254, App
Sequence 259, App
Sequence 260, App
Sequence 241, App
Sequence 3, Appli
Sequence 20, Appl
Sequence 29, Appl

6 US-10-732-923-4727
6 US-10-732-923-13054
6 US-10-732-923-18303
6 US-10-915-740A-1047
6 US-10-732-923-19147
6 US-10-732-923-19146
6 US-10-511-632-1
6 US-10-732-923-19212
6 US-10-837-269-265
6 US-10-837-269-271
6 US-10-837-269-277
6 US-10-837-269-283
6 US-10-837-269-253
6 US-10-837-269-254
6 US-10-837-269-259
6 US-10-837-269-260
6 US-10-837-269-241
6 US-10-837-269-3
6 US-10-837-269-20
6 US-10-837-269-29

60.5 12.4 667
60.5 12.4 682
60.5 12.4 1330
60.5 12.4 1457
60 12.3 2307
60 12.3 2314
59.5 12.2 365
59 12.1 303
59 12.1 1009
59 12.1 1009
59 12.1 1009
59 12.1 1011
59 12.1 1011
59 12.1 1011
59 12.1 1017
59 12.1 1021
59 12.1 1021
59 12.1 1021

```
; PRIOR APPLICATION NUMBER: 60/460,103
; PRIOR FILING DATE: 2003-04-02
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 105
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(51)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-805-684-105

Query Match          13.9%; Score 67.5; DB 6; Length 679;
Best Local Similarity 35.1%; Pred. No. 7.4; 27; Indels 11; Gaps 4;
Matches 26; Conservative 10; Mismatches 27; Indels 11; Gaps 4;

QY 1 TLTPAYIDPENVKKTGNV--TLATIGS-----TTANKTVLDVDEYKYGKFNLSIKGTIDQ 55
Db 348 TSTQPAAL-----VSSGSASTLASAINADLFATASADVAADIATYTSKMDAIGKGTWE 403
;
QY 56 I--DKTNNVTYQTI 67
Db 404 IYNDLSKNTTGSTI 417
;
RESULT 3
US-10-837-269-266
; Sequence 266, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: 109P1D4 Useful in Treatment and Corresponding Proteins Entitled
; FILE REFERENCE: 51158-20083.00
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 266
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-266

Query Match          13.1%; Score 64; DB 6; Length 1009;
Best Local Similarity 27.8%; Pred. No. 29; 43; Indels 22; Gaps 5;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTPAYIDPENVKKTGNVTLA--TGIGSTTANKTVLVD-----YEKYGKFY-- 45
Db 520 LTVVKLDREKEDKYLEILAKDNGVPPVTSNVTVFVSIIDQNDNSPVFTNEY-KFYVP 578
;
QY 46 -NLSIKGTIDQIDKTNNTYRQ---TIVNPSGDNVIAPVLGTGNLKP 88
Db 579 ENLPRHGTGVLITVDPDYGDNSAVTSLSDENDDDFTIDSGTGVIRPN 626
;
RESULT 5
US-10-837-269-278
; Sequence 278, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 51158-20083.00
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-278

Query Match          13.1%; Score 64; DB 6; Length 1009;
Best Local Similarity 27.8%; Pred. No. 29; 43; Indels 22; Gaps 5;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTPAYIDPENVKKTGNVTLA--TGIGSTTANKTVLVD-----YEKYGKFY-- 45
; GENERAL INFORMATION:
```

```
Db 520 LTVVKLDREKDKYLFITLAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 578
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPNGDNVIAPIVLTGNLKP 88
Db 579 ENLPRHGTGLITVTPDYGNSAVTSLILDENDDFIDSQTGVIRPN 626

RESULT 6
US-10-837-269-284
; Sequence 284, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 109PID4 Useful in Treatment and Detection of Cancer
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-284

Query Match 13.1%; Score 64; DB 6; Length 1009;
Best Local Similarity 27.8%; Pred. No. 29;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPIVDENVKKTGNVTIA--TGIGSTANKTVLVD-----YEYKGFY-- 45
Db 520 LTVVKLDREKDKYLFITLAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 578
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPNGDNVIAPIVLTGNLKP 88
Db 579 ENLPRHGTGLITVTPDYGNSAVTSLILDENDDFIDSQTGVIRPN 626

RESULT 7
US-10-837-269-13
; Sequence 13, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 109PID4 Useful in Treatment and Detection of Cancer
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-13

Query Match 13.1%; Score 64; DB 6; Length 1037;
Best Local Similarity 27.8%; Pred. No. 30;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPIVDENVKKTGNVTIA--TGIGSTANKTVLVD-----YEYKGFY-- 45
Db 520 LTVVKLDREKDKYLFITLAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 578
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPNGDNVIAPIVLTGNLKP 88
Db 579 ENLPRHGTGLITVTPDYGNSAVTSLILDENDDFIDSQTGVIRPN 626

RESULT 8
US-10-837-269-19
; Sequence 19, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 109PID4 Useful in Treatment and Detection of Cancer
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-19

Query Match 13.1%; Score 64; DB 6; Length 1037;
Best Local Similarity 27.8%; Pred. No. 30;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPIVDENVKKTGNVTIA--TGIGSTANKTVLVD-----YEYKGFY-- 45
Db 543 LTVVKLDREKDKYLFITLAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 601
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPNGDNVIAPIVLTGNLKP 88
Db 602 ENLPRHGTGLITVTPDYGNSAVTSLILDENDDFIDSQTGVIRPN 649

RESULT 9
US-10-837-269-25
; Sequence 25, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 109PID4 Useful in Treatment and Detection of Cancer
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-25

Query Match 13.1%; Score 64; DB 6; Length 1037;
Best Local Similarity 27.8%; Pred. No. 30;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPIVDENVKKTGNVTIA--TGIGSTANKTVLVD-----YEYKGFY-- 45
Db 543 LTVVKLDREKDKYLFITLAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 601
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPNGDNVIAPIVLTGNLKP 88
Db 602 ENLPRHGTGLITVTPDYGNSAVTSLILDENDDFIDSQTGVIRPN 649
```

```
; Sequence 36, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; TITLE OF INVENTION: 109PID4 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20083.00
; CURRENT APPLICATION NUMBER: US/10/837,269
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-837-269-36

Query Match      13.1%; Score 64; DB 6; Length 1037;
Best Local Similarity 27.8%; Pred. No. 30;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

Qy      2 LTMPAYIDPENVKKTGNVTLA--TGIGSTTANKTVLVD-----YEKYGKFY-- 45
Db      543 LTVVKLDREKEDKYLFTILAKDNGVPPLTSNTVFSIIQNDNSPVFTTNEY-KFYVP 601

Qy      46 -NLSIKGTIDQIDKTNNTYRQ----TIYVNPSPGDNVIAPLVLTGNLKP 88
Db      602 ENLPRHGTGVLITVTDPDYGDNSAVTLSILDENDDFTIDSQTGVIRPN 649

RESULT 12
US-10-837-269-39
; Sequence 39, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; TITLE OF INVENTION: 109PID4 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20083.00
; CURRENT APPLICATION NUMBER: US/10/837,269
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-837-269-39

Query Match      13.1%; Score 64; DB 6; Length 1037;
Best Local Similarity 27.8%; Pred. No. 30;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

Qy      2 LTMPAYIDPENVKKTGNVTLA--TGIGSTTANKTVLVD-----YEKYGKFY-- 45
Db      543 LTVVKLDREKEDKYLFTILAKDNGVPPLTSNTVFSIIQNDNSPVFTTNEY-KFYVP 601

Qy      46 -NLSIKGTIDQIDKTNNTYRQ----TIYVNPSPGDNVIAPLVLTGNLKP 88
Db      602 ENLPRHGTGVLITVTDPDYGDNSAVTLSILDENDDFTIDSQTGVIRPN 649

RESULT 11
US-10-837-269-36
; Sequence 28, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; TITLE OF INVENTION: 109PID4 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20083.00
; CURRENT APPLICATION NUMBER: US/10/837,269
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-837-269-28

Query Match      13.1%; Score 64; DB 6; Length 1037;
Best Local Similarity 27.8%; Pred. No. 30;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

Qy      2 LTMPAYIDPENVKKTGNVTLA--TGIGSTTANKTVLVD-----YEKYGKFY-- 45
Db      543 LTVVKLDREKEDKYLFTILAKDNGVPPLTSNTVFSIIQNDNSPVFTTNEY-KFYVP 601

Qy      46 -NLSIKGTIDQIDKTNNTYRQ----TIYVNPSPGDNVIAPLVLTGNLKP 88
Db      602 ENLPRHGTGVLITVTDPDYGDNSAVTLSILDENDDFTIDSQTGVIRPN 649

RESULT 11
US-10-837-269-36
```

```

Qy      2  LTMFAYIDPENVKKTGNVTLA--TGIGSTTANKTVLVD-----YEKYKIFY-- 45
Db      543  LTVVKKLDREKEDKYLFTILKNDGVPPLTSNVTVFVSIIDQNDSPVFTHNEY-KFYVP 601

Qy      46  -NLSIKGTIDQDKTNTYRQ----TIYVNSGDNVIAPVLTGNLKP 88
Db      602  ENLPRHGTGVGII TVDPDYGNSAVTLGIIDNDQDFTIDSGTVIRPN 649

```

```

RESULT 13
US-10-837-269-264
; Sequence 264, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; "109P1d4 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20083.00
; CURRENT APPLICATION NUMBER: US/10/837,269
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-264

```

RESULT 14
US-10-837-269-282
/ Sequence 282, Application US/10837269
/ GENERAL INFORMATION:
/ APPLICANT: Raitano, Arthur B.
/ APPLICANT: Challita-Rid, Pia M.
/ APPLICANT: Ge, Wangmao
/ APPLICANT: Perez-Villar, Juan J.
/ APPLICANT: Kanner, Steven B.
/ APPLICANT: Jakobovits, Aya
/ TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
/ TITLE OF INVENTION: 109PID4 Useful in Treatment and Detection of Cancer
/ FILE REFERENCE: 51158-20083.00
/ CURRENT APPLICATION NUMBER: US/10/837,269
/ CURRENT FILING DATE: 2004-04-30
/ PRIOR APPLICATION NUMBER: US10/121,024
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: US 60/467,002
/ PRIOR FILING DATE: 2003-04-30
/ PRIOR APPLICATION NUMBER: US 60/282,739
/ PRIOR FILING DATE: 2001-04-10
/ NUMBER OF SEQ ID NOS: 284

Search completed: November 3, 2004, 19:30:35
Job time : 2.37768 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:53:31 ; Search time 19.4464 Seconds
(without alignments)
2781.239 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_425

Perfect score: 487

Sequence: 1 TLTPAVIDPENVKKTGNVT.....NVIAPVLGNLKPNTDSNAL 94

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	100.0	933	2	Q53653
2	481	98.8	1029	2	Q6GIK4
3	470	96.5	928	2	Q6GB45
4	470	96.5	946	2	Q8NXX1
5	422	86.7	935	2	Q932C5
6	422	86.7	989	2	Q93VJ4
7	329	67.6	881	2	Q93MH7
8	113	23.2	741	2	Q6EVM1
9	113	23.2	1018	1	FNBA_STAAT
10	107.5	22.1	965	2	Q6GDJ5
11	102.5	21.0	931	2	Q9KI13
12	100.5	20.6	677	2	Q6EVM0
13	100.5	20.6	940	2	Q53682
14	99.5	20.4	1056	2	Q8CQ72
15	99.5	20.4	1092	2	Q70022
16	99.5	20.4	1137	2	Q6GJA6
17	99.5	20.4	1171	2	Q9KXW6
18	90.5	18.6	943	2	Q8NUU8
19	90.5	18.6	957	2	Q6G6H4
20	88.5	18.2	1141	2	Q6GBS4
21	88.5	18.2	1141	2	Q8NXX5
22	88.5	18.2	1166	2	Q8G489
23	87.5	18.0	961	2	Q99RD3
24	87.5	18.0	961	2	Q7A3J8
25	87	17.9	1141	2	Q99W46
26	87	17.9	1141	2	Q932F7
27	84.5	17.4	1015	2	Q6G6H3
28	84.5	17.4	1015	2	Q8NUU7
29	84.5	17.4	1038	2	Q99RD2
30	84.5	17.4	1038	2	Q7A3J7
31	82	16.8	434	2	Q7N3E9

32	81.5	16.7	1315	2	Q86488	staphylococ
33	81.5	16.7	1365	2	Q6GBS5	staphylococ
34	81	16.6	1893	2	Q8KWM1	staphylococ
35	80.5	16.5	444	2	Q8EUV6	mycoplasma
36	80.5	16.5	1347	2	Q8NXX6	staphylococ
37	79	16.2	1633	2	Q8CMP4	staphylococ
38	79	16.2	1733	2	Q9K114	staphylococ
39	78	16.0	202	1	DR3E_AQVAB	staphylococ
40	78	16.0	1129	2	Q8XLL1	staphylococ
41	77	15.8	856	1	ALE7_AZOV1	clostridium
42	76.5	15.7	1183	1	YNK6_YEAST	azotobacter
43	76.5	15.7	2170	2	Q81DI8	saccharomyc
44	76	15.6	898	2	Q86JG0	plasmodium
45	76	15.6	1295	2	Q93SV6	dictyosteli
						Q93SV6 chlorobium

ALIGNMENTS

RESULT 1
Q53653 ID Q53653 PRELIMINARY; PRT; 933 AA.
AC Q53653;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Clumping factor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RX MEDLINE=94224142; PubMed=8170386;
RA McDewitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor (fibrogen receptor
of Staphylococcus aureus";
RL Mol. Microbiol. 11:237-248(1994).
CC -! SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
DR EMBL; Z18852; CAA79304.1; -.
DR PIR; S41539; S41539.
DR PDB; 1N67; X-ray; A=202-560.
DR GO; GO:0009886; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 933 AA; 97057 MW; EB51A6DE2FF759F4 CRC64;

Query Match 100.0%; Score 487; DB 2; Length 933;
Best Local Similarity 100.0%; Pred. No. 6.2e-39;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TLTPAYIDPENVKKTGNVTATGIGSTANKTVLVDYKYGKFNLSIKGTIDQIDKTN	60
Db	332	TLTPAYIDPENVKKTGNVTATGIGSTANKTVLVDYKYGKFNLSIKGTIDQIDKTN	391
QY	61	NTYRQTIYVNSGDNVIAPIVLTGNLKPNTDSNAL	94
Db	392	NTYRQTIYVNSGDNVIAPIVLTGNLKPNTDSNAL	425

RESULT 2
Q6GIK4 ID Q6GIK4 PRELIMINARY; PRT; 1029 AA.
AC Q6GIK4;
DT 05-JUL-2004 (Tremblrel. 27, Created)

```

RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -i- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; BX571857; CAG42526.1; -.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Peptidoglycan-anchor.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 928 AA; 96417 MW; FBE8570209B9E195 CRC64;

Query Match 96.5%; Score 470; DB 2; Length 928;
Best Local Similarity 96.8%; Pred. No. 2.8e-37;
Matches 90; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LTMPAYIDPENVKKTGNVTATGIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTN 61
Db 332 ITMPAYIDPENVTGTGNVTLTGIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTN 391
QY 62 TYRTIYVPSGDNVIAPIVLTGNLKPNTDSNAL 94
Db 392 TYRTIYVPSGDNVIAPIVLTGNLKPNTDSNAL 424

RESULT 4
Q8NXXJ1 PRELIMINARY; PRT; 946 AA.
AC Q8NXXJ1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein.
GE Name=clfa; OrderedLocusNames=MW0764;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12043378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -i- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AF004824; BAB94629.1; -.
DR HSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 946 AA; 98237 MW; EFPB83793201173 CRC64;

Query Match 96.5%; Score 470; DB 2; Length 946;
Best Local Similarity 96.8%; Pred. No. 2.9e-37;
Matches 90; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LTMPAYIDPENVKKTGNVTATGIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTN 61
Db 332 ITMPAYIDPENVTGTGNVTLTGIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTN 391

```

```
QY 62 TYRQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNAL 94
Db 392 TYRQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNAL 424

RESULT 5
Q932C5 PRELIMINARY; PRT; 935 AA.
AC Q932C5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein.
GN Name=fnb; OrderedLocusNames=SAV0811;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699;
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RN Lancel 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
DR EMBL; AP003360; BAB56973.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor.
DR Pfam; PF04650; Y5IRK_signal.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 935 AA; 96949 MW; DC5A2D92CE3BA91C CRC64;

Query Match 86.7%; Score 422; DB 2; Length 935;
Best Local Similarity 84.9%; Pred. No. 1.5e-32;
Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 LTPAYIDPENVKKTGNVLTATGIGSTANKTVLVDYKYGKYNLSIKGTIDQIDKTN 61
Db 333 LTPAYIDPENVKKTGNVLTATGIGSTANKTVLVDYKYGKYNLSIKGTIDQIDKTN 392

RESULT 6
Q99VJ4 PRELIMINARY; PRT; 989 AA.
AC Q99VJ4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein A, clumping factor.
GN Name=clfA; OrderedLocusNames=SA0742;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N315;
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RN Lancel 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
DR EMBL; AP003360; BAB56973.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor.
DR Pfam; PF04650; Y5IRK_signal.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 935 AA; 96949 MW; DC5A2D92CE3BA91C CRC64;

Query Match 86.7%; Score 422; DB 2; Length 935;
Best Local Similarity 84.9%; Pred. No. 1.5e-32;
Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 LTPAYIDPENVKKTGNVLTATGIGSTANKTVLVDYKYGKYNLSIKGTIDQIDKTN 61
Db 333 LTPAYIDPENVKKTGNVLTATGIGSTANKTVLVDYKYGKYNLSIKGTIDQIDKTN 392

RESULT 7
Q93MH7 PRELIMINARY; PRT; 881 AA.
AC Q93MH7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Receptor Fbl precursor.
GN Name=fbl;
OS Staphylococcus lugdunensis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=28035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2342;
RX Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
DR EMBL; AF04823; AAK95649.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK_signal; 1.
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=N315;
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RN Lancel 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
DR EMBL; AP003311; BAB41975.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;

Query Match 86.7%; Score 422; DB 2; Length 989;
Best Local Similarity 84.9%; Pred. No. 1.6e-32;
Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 LTPAYIDPENVKKTGNVLTATGIGSTANKTVLVDYKYGKYNLSIKGTIDQIDKTN 61
Db 333 LTPAYIDPENVKKTGNVLTATGIGSTANKTVLVDYKYGKYNLSIKGTIDQIDKTN 392

RESULT 7
Q93MH7 PRELIMINARY; PRT; 881 AA.
AC Q93MH7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Receptor Fbl precursor.
GN Name=fbl;
OS Staphylococcus lugdunensis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=28035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2342;
RX Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
DR EMBL; AF04823; AAK95649.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK_signal; 1.
```

```
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchored; Signal.
FT SIGNAL 1 39 Potential.
FT CHAIN 40 881
SQ SEQUENCE 881 AA; 94250 MW; D4296C4959C4F19B CRC64;

Query Match 67.6%; Score 329; DB 2; Length 881;
Best Local Similarity 66.7%; Pred. No. 1.8e-23;
Matches 62; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 2 LTMPAYIDPENVKKTGNVLTATIGSTTANKVLVDYKYGKFNLSIKGTIDQIDKTN 61
Db 317 ISIPGYIDPNVHTGKVNLTSGITQAKTIVVDYKYGKFNLSIKGTIDQIDKVN 376
QY 62 TYRTQTVYNPSGDNVIAPLTGNLKPNTDSN 94
Db 377 TYRTQTVYNPSGDNVIAPLTGNLKPNTDSN 409

RESULT 8
Q6EVM1 PRELIMINARY; PRT; 741 AA.
ID AC Q6EVM1
AC Q6EVM1;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fibronectin binding protein A precursor.
GN Name=fnbA;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman D2C;
RA Grundmeier M., Hussain M., Becker P., Sinha B.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ629121; CAF32747.1; --
KW Signal.
FT SIGNAL 1 36 Potential.
FT CHAIN 37 741 fibronectin binding protein A.
FT SEQUENCE 741 AA; 81433 MW; 54D9A417128AE7EC CRC64;
SQ SEQUENCE 741 AA; 81433 MW; 54D9A417128AE7EC CRC64;

Query Match 23.2%; Score 113; DB 2; Length 741;
Best Local Similarity 25.0%; Pred. No. 0.023;
Matches 23; Conservative 24; Mismatches 43; Indels 2; Gaps 2;

QY 2 LTMPAYIDPENVKKTGNVLTATIGSTTANKVLVDY-EYKGFYNLSIKGTIDQIDKTN 60
Db 301 LEINLFDPKVTQTNNGNQTITSLNEEQTSKELDVYKDGIGNY-ANLNGSIETFNKAN 359
QY 61 NTRYQTVYNPSGDNVIAPLTGNLKPNTDSN 92
Db 360 NRFSHVAFIKPNNGKTTSTVTGTLMKGSNQ 391

RESULT 9
FNBA_STAAU STANDARD; PRT; 1018 AA.
ID AC FNBA_STAAU
AC FNBA_STAAU;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Fibronectin-binding protein precursor (FNBP).
GN Name=fnbA;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 8325-4;
RX MEDLINE=89098998; PubMed=25213191;
RA Signaes C., Raucchi G., Joensson K., Lindgren P.-E.,
RA Anantharamaiah G.M., Hoeck M., Lindberg M.;

"Nucleotide sequence of the gene for a fibronectin-binding protein from Staphylococcus aureus: use of this peptide sequence in the synthesis of biologically active peptides.";
Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).
-1- FUNCTION: The ability of bacteria to bind fibronectin has been proposed as a virulence factor enabling bacteria to colonize wound tissues and blood clots. Binding of plasma fibronectin to the bacterial surface might block adhesion receptors on S.aureus, thus representing an important defense mechanism against tissue invasion.
-1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (Potential).
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; J04151; AAA26632.1; --
InterPro; IPR004237; Fn_bind.
InterPro; IPR005877; SPos_YSIK.
InterPro; IPR001899; Gram_pos_anchor.
Pfam; PF02986; Fn_bind; 1.
Pfam; PF00746; Gram_pos_anchor; 1.
Pfam; PF04650; YSIK_signal; 1.
TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
TIGRFAMs; TIGR01168; YSIK_signal; 1.
PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell adhesion; Cell wall; Peptidoglycan-anchored; Repeat; Signal; Virulence.
FT SIGNAL 1 36
FT CHAIN 37 985 Fibronectin-binding protein.
FT PROPEP 986 1018 Removed by sortase (Potential).
FT REPEAT 545 574 B-1.
FT REPEAT 575 604 B-2.
FT REPEAT 745 878 4 X APPROXIMATE TANDEM REPEATS,
FT DOMAIN 745 878 FIBRONECTIN-BINDING DOMAIN.
FT REPEAT 745 782 D-1.
FT REPEAT 783 820 D-2.
FT REPEAT 821 859 D-3.
FT REPEAT 860 878 D-4 (incomplete).
FT DOMAIN 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).
FT REPEAT 893 906 WR 1.
FT REPEAT 907 920 WR 2.
FT REPEAT 921 934 WR 3.
FT REPEAT 935 948 WR 4.
FT REPEAT 948 985 WR 5.
FT SITE 985 985 LPXTG sorting signal (Potential).
FT MOD_RES 985 985 Pentaglycyl murein peptidoglycan amidated threonine (Potential).
FT SEQUENCE 1018 AA; 111780 MW; 58175E0020B81F1F CRC64;
SQ SEQUENCE 1018 AA; 111780 MW; 58175E0020B81F1F CRC64;

Query Match 23.2%; Score 113; DB 1; Length 1018;
Best Local Similarity 25.0%; Pred. No. 0.032;
Matches 23; Conservative 24; Mismatches 43; Indels 2; Gaps 2;

QY 2 LTMPAYIDPENVKKTGNVLTATIGSTTANKVLVDY-EYKGFYNLSIKGTIDQIDKTN 60
Db 301 LEINLFDPKVTQTNNGNQTITSLNEEQTSKELDVYKDGIGNY-ANLNGSIETFNKAN 359
QY 61 NTRYQTVYNPSGDNVIAPLTGNLKPNTDSN 92
Db 360 NRFSHVAFIKPNNGKTTSTVTGTLMKGSNQ 391

RESULT 10
Q6GDU5 PRELIMINARY; PRT; 965 AA.
ID AC Q6GDU5
AC Q6GDU5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
```

05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
Name=fba; ORFNames=SAR2580;
Staphylococcus aureus subsp. aureus MRSA252.
Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=282458;
[1]
SEQUENCE FROM N.A.
RC STRAIN=MRSA252;
RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltham T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
EMBL: BX571856; CAG41560.1;
DR InterPro: IPR04237; Fb_bind.
DR InterPro: IPR005877; Gpos_YsIRK.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF02986; Fb_bind; 1.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF04650; YsIRK_signal; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs: TIGR01168; YsIRK_signal; 1.
DR PROSITE: PS00847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Peptidoglycan-anchor.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 965 AA; 105691 MW; 2982E132164D0551 CRC64;

Query Match 22.1%; Score 107.5; DB 2; Length 965;
Best Local Similarity 25.0%; Pred. No. 0.11;
Matches 23; Conservative 23; Mismatches 45; Indels 1; Gaps 1;

QY 2 LMPAYIDPENVKKTGNVLTATGIGSTTANKTVLVVDYKYKFNLSIKGTIDQDKTN 61
Db LEINLFDPTQSQNQQTITKNGKSTGTMQITKDGVRNQXTNVNGSIETFDKKN 360

QY 62 TYRQTIYVNP-SGDNVIAPLVTGNLKPNTDSN 92
Db KTHVAYIKPINGSNDSVTVTGMLTQGSNEN 392

RESULT 11
Q9KIL3
ID Q9KIL3 PRELIMINARY; PRT; 931 AA.
AC Q9KIL3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fibrinogen-binding protein SdrG.
GN Name=sdrG;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K28;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RT epidermidis."
RL Microbiology 146:1535-1546(2000).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).

EMBL: AF245042; AAF72510.1;
DR HSSP: Q53653; IN67.
DR GO: GO:0009966; C:cell surface; IEA.
DR GO: GO:0005618; C:cell wall; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR008454; Cna_B.
DR InterPro: IPR005877; Gpos_YsIRK.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF05738; Cna_B; 2.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF04650; YsIRK_signal; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs: TIGR01168; YsIRK_signal; 1.
DR PROSITE: PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 931 AA; 102955 MW; 591B657D97027116 CRC64;

Query Match 21.0%; Score 102.5; DB 2; Length 931;
Best Local Similarity 32.9%; Pred. No. 0.31;
Matches 24; Conservative 13; Mismatches 31; Indels 5; Gaps 2;

QY 2 LMPAYIDPENVKKTG---NVLTATGIGSTTANKTVLVVDYKYKFNLSIKGTIDQDK 58
Db LKLTSTYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQKPNENRTANLQSMFTNIDT 444

QY 59 TNNYRQTIYVNP 71
Db KHTVEQTIYINP 457

RESULT 12
Q6EVM0
ID Q6EVM0 PRELIMINARY; PRT; 677 AA.
AC Q6EVM0;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fibrinectin binding protein B precursor.
GN Name=fnbB;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman D2C;
RA Grundmeier M., Hussain M., Becker P., Sinha B.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ629122; CAF32748.1;
KW Signal.
FT SIGNAL 1 35 Potential.
FT CHAIN 35 677 fibronectin binding protein B.
SQ SEQUENCE 677 AA; 74737 MW; 37CAD12EFD160B1 CRC64;

Query Match 20.6%; Score 100.5; DB 2; Length 677;
Best Local Similarity 23.2%; Pred. No. 0.34;
Matches 23; Conservative 25; Mismatches 38; Indels 13; Gaps 3;

QY 2 LMPAYIDPENVKKTGNVLTATGIGSTTANKTVLVVDY-----EKYKFNLSIKGTIDQI 56
Db LSLNLFIDPTTQKGNQNVKLGGETTVSKIFNIQYLGVRDNGW----VTANGRIDTL 328

QY 57 DKTNTYRQTIYVNPSPGDNVIAPLVTGNL-----KPNTDS 91
Db NKVDGKFSHFAYMKPNNSQSLSSVTVTGQTKNGKPGVNN 367

RESULT 13
Q53682
ID Q53682 PRELIMINARY; PRT; 940 AA.
AC Q53682;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```

CC      an amide bond (By similarity).
DR      EMBL: AE016745; AAC03928.1; -.
DR      HSSP: Q53653; 1N67.
DR      GO: GO:0009986; C:cell surface; IEA.
DR      GO: GO:0005618; C:cell wall; IEA.
DR      GO: GO:0016020; C:membrane; IEA.
DR      InterPro: IPR008454; Cna_B.
DR      InterPro: IPR005877; Gpos_Ysirk.
DR      InterPro: IPR001899; Gram_pos_anchor.
DR      InterPro: IPR008964; Invasin_intimin.
DR      Pfam: PF05738; Cna_B; 2.
DR      Pfam: PF00746; Gram_pos_anchor; 1.
DR      Pfam: PF04650; Ysirk_signal; 1.
DR      TIGRFAMS: TIGR01167; LPXTG_anchor; 1.
DR      TIGRFAMS: TIGR01168; Ysirk_signal; 1.
DR      PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW      Cell wall; Complete proteome; Peptidoglycan-anchor.
KW      SEQUENCE 1055 AA; 115727 MW; 015869A9B5CA2723 CRC64;
SQ

Query Match      20.4%; Score 99.5; DB 2; Length 1056;
Best Local Similarity 31.5%; Pred.No.0.71; Indels 5; Gaps 2;
Matches 23; Conservative 14; Mismatches 31;

QY      2 LTPAYIDPNVKKTG---NVLTATIGSTANKTLVLDYKYGKFNLSIKGTIDQDK 58
DB      389 LKLTSYIDKSKVNNNTKLDVEYKTLSS--VNKTIIVYQRPNEKRTANLQSMFTNIDT 446
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      59 TNNYRQTIYVNP 71
DB      447 KNHTVEQTIYVNP 459
      - - - - - - - - - - - - - - - - - - - - - - - - - - - -

RESULT 15
O70022      PRELIMINARY; PRT; 1092 AA.
AC      O70022;
DT      01-AUG-1998 (TrEMBLrel. 07, Created)
DT      01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Fibrinogen-binding protein precursor.
OS      Staphylococcus epidermidis.
OC      Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX      NCBI_TaxID=1282;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=HB;
RX      MEDLINE=98261511; PubMed=9596732;
RA      Nilsson M., Frykberg L., Flock J.I., Pei L., Lindberg M., Guss B.;
RT      "A fibrinogen-binding protein of Staphylococcus epidermidis.";
RL      Infect. Immun. 66:2666-2673(1998).
CC      -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC      an amide bond (By similarity).
CC      EMBL: Y17116; CAA76638.1; -.
DR      PIR: T30214; T30214.
DR      HSSP: Q53653; 1N67.
DR      GO: GO:0009986; C:cell surface; IEA.
DR      GO: GO:0005618; C:cell wall; IEA.
DR      GO: GO:0016020; C:membrane; IEA.
DR      InterPro: IPR008454; Cna_B.
DR      InterPro: IPR005877; Gpos_Ysirk.
DR      InterPro: IPR001899; Gram_pos_anchor.
DR      InterPro: IPR008964; Invasin_intimin.
DR      Pfam: PF05738; Cna_B; 2.
DR      Pfam: PF00746; Gram_pos_anchor; 1.
DR      Pfam: PF04650; Ysirk_signal; 1.
DR      TIGRFAMS: TIGR01167; LPXTG_anchor; 1.
DR      TIGRFAMS: TIGR01168; Ysirk_signal; 1.
DR      PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW      Cell wall; Peptidoglycan-anchor; Signal.
KW      SIGNAL 1 51 Potential.
FT      CHAIN 52 1092 Fibrinogen-binding protein.
SQ      SEQUENCE 1092 AA; 119292 MW; 6542BC39AAD8B984 CRC64;

```

```

Query Match      20.4%; Score 99.5; DB 2; Length 1092;
Best Local Similarity 31.5%; Pred. No. 0.74;
Matches 23; Conservative 14; Mismatches 31; Indels 5; Gaps 2;

QY 2 LTMPEYIDPENVKKTG--NVTLATGIGSTTANKTVLVDYEKYGKFFYNLSIKGTIDQIDK 58
   | : : : | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 389 LKLTSYIDKSKVPNNNTKLDVEYKTAISS--VNKTIVVEYORPNNRTANLQSMFTNIDT 446

QY 59 TNNTYRQTIYVNP 71
   | : | | | | : |
Db 447 KNHTVEQTIYINP 459

```

Search completed: November 3, 2004, 19:10:19
Job time : 20.4464 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:48:10 ; Search time 41.4791 Seconds
(without alignments)
1894.005 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_550

Perfect score: 1157

Sequence: 1 TLTPAYIDPENVKYTGNTV.....NGSGSGDGDKPVVPEQPD 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003s.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1157	100.0	331	5	Aae29263 Staphyloc
2	1157	100.0	345	2	Aaw31555 Fibronect
3	1157	100.0	520	5	Aae29262 Staphyloc
4	1157	100.0	927	6	Abm72221 Staphyloc
5	1157	100.0	933	3	Aay58435 Staphyloc
6	1157	100.0	933	4	Aab69508 Staphyloc
7	1157	100.0	933	6	Abj18947 Pathogen
8	1157	100.0	936	2	Aaw89801 Staphyloc
9	991	85.7	935	6	Abu16402 Protein e
10	991	85.7	1021	4	Aau33975 Staphyloc
11	991	85.7	1021	4	Aau36951 Staphyloc
12	863	74.6	496	5	Aau75490 S. aureus
13	863	74.6	496	6	Ada89664 Staphyloc
14	863	74.6	496	6	Ada89663 Staphyloc
15	255	22.0	767	4	Aau34403 Staphyloc
16	255	22.0	767	4	Aau37547 Staphyloc
17	255	22.0	940	2	Aar07070 Fibronect
18	255	22.0	940	6	Abj19111 Pathogen
19	255	22.0	948	6	Ada89470 Staphyloc
20	255	22.0	948	6	Abj18923 Pathogen
21	255	22.0	948	6	Abm72536 Staphyloc
22	244.5	21.1	343	7	Abm79016 Staphyloc
23	244.5	21.1	560	7	Abm79015 Staphyloc
24	244.5	21.1	892	6	Abu42557 Protein e
25	244.5	21.1	892	8	Ada84849 S epiderm

26	244.5	21.1	930	5	ABP40469
27	244.5	21.1	991	3	AY83171 Cell wall
28	244.5	21.1	991	3	AY70120 Staph. ep
29	241.5	20.9	549	7	ABM79020 Staphyloc
30	241.5	20.9	1092	2	AAW41602 Staphyloc
31	241.5	20.9	1092	7	ABM79019 Staphyloc
32	237.5	20.5	961	6	ABU15854 Protein e
33	208	18.0	567	6	ADA89555 Staphyloc
34	208	18.0	743	6	ADA89690 Staphyloc
35	208	18.0	877	6	ADA89539 Staphyloc
36	208	18.0	877	6	ABU42504 Protein e
37	208	18.0	877	6	ABM72702 Staphyloc
38	208	18.0	913	6	ABJ18917 Pathogen
39	208	18.0	918	2	AY88640 S. aureus
40	207.5	17.9	1166	2	AY88643 S. aureus
41	207.5	17.9	1166	6	ABJ18982 Pathogen
42	200.5	17.3	1141	6	ABU42327 Protein e
43	196	16.9	316	7	ABM79017 Staphyloc
44	187.5	16.2	670	6	ABU42520 Protein e
45	183.5	15.9	1018	4	Aau37245 Staphyloc

ALIGNMENTS

RESULT 1

AAE29263
ID AAE29263 standard; protein; 331 AA.

XX AAE29263;

XX AC

XX 27-JAN-2003 (first entry)

XX DT

XX XX

DE Staphylococcus aureus Clf33 protein.

XX XX

KW Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein;
immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.

XX OS

XX Staphylococcus aureus.

XX XX

XX WO200273600-A2.

XX XX

XX PD

XX 19-SEP-2002.

XX XX

XX PF

XX 28-JAN-2002; 2002WO-US002296.

XX PR

XX 26-JAN-2001; 2001US-0264072P.

XX PR

XX 12-MAR-2001; 2001US-0274611P.

XX PR

XX 18-JUN-2001; 2001US-0298413P.

XX PR

XX 30-JUL-2001; 2001US-0308116P.

XX PA

XX (INHI-) INHIBITEX INC.

XX XX

XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;

XX WPI; 2002-759334-02.

XX N-PSDB; AAD46862.

XX XX

XX New anti- clumping factor A (ClfA) monoclonal antibody, useful for

XX treating or preventing Staphylococcus aureus infection e.g. wound

XX infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in

XX a human or animal.

XX Claim 9; Page 70-72; 80pp; English.

XX XX

XX The invention relates to monoclonal antibody which binds the clumping

XX factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA

XX monoclonal antibody is useful for treating or preventing S. aureus

XX infection in a human or animal, and for inhibiting the binding of

XX staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment

XX S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3

XX protein is useful for inducing an immunological response in a human or

CC animal. These staphylococcal infections include wound infections, sepsis,
CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC present sequence is Staphylococcus aureus Clf33 protein
XX
XX
SQ Sequence 331 AA;

Query Match 100.0%; Score 1157; DB 5; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.le-86;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLMPAYIDPENVKKTGNVTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
DB 113 TLMPAYIDPENVKKTGNVTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 172

QY 61 NTRYQTIVNPSGDNVIAPVLTGNLKTNTSNALIDQONTSIKVKYVDNAADLSSEYFVN 120
DB 173 NTRYQTIVNPSGDNVIAPVLTGNLKTNTSNALIDQONTSIKVKYVDNAADLSSEYFVN 232

QY 121 PENFEDVTNSVNIITFPNPNQYKVFETPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 180
DB 233 PENFEDVTNSVNIITFPNPNQYKVFETPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 292

QY 181 YNSNIIWRSMWDNEVAFNNGSGGSGDGIDKVPVPEQDPE 219
DB 293 YNSNIIWRSMWDNEVAFNNGSGGSGDGIDKVPVPEQDPE 331

RESULT 2
AAW31555
ID AAW31555 standard; protein; 345 AA.
XX
XX AAW31555;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 21-MAY-1998 (first entry)
XX
XX
XX Fibronectin-binding MSCRAMM derivative pCF33.
XX
XX Fibronectin; pCF33; collagen binding protein; sepsis; infection;
KW microbial surface component recognising adhesive matrix molecule; MSCRAMM;
KW adhesin; vaccine; immunisation; diagnosis; therapy; mastitis.
XX
XX Staphylococcus aureus.
XX
XX
XX Key Location/Qualifiers
FH Peptide 1..12
FT /note= "vector pQE30-derived peptide"
XX
XX WO9743314-A2.
XX
XX 20-NOV-1997.
XX
XX 14-MAY-1997; 97WO-US008210.
XX
XX 16-MAY-1996; 96US-0017678P.
XX (TEXA) UNIV TEXAS A & M SYSTEM.
XX (UABR-) UAB-RES FOUND.
XX
XX Hoeek M, Patti JM, House-Pompeo K, Sthanam N, Symersky J;
XX WPI; 1998-008601/01.
XX
XX Antibody that interacts with collagen binding domain of Staphylococcal
PT cna gene product - useful to prevent bacterial sepsis in animal infected
PT with Staphylococcus aureus.
XX
XX Disclosure; Page 91; 143pp; English.
XX
XX This protein comprises Staphylococcus aureus fibronectin-binding
CC microbial surface component recognising adhesive matrix molecule (MSCRAMM)
CC derivative pCF33, plus a vector-derived N-terminal peptide. The invention

CC relates to claimed nucleic acid sequences (see AAT93436-38) encoding S.
CC aureus collagen binding protein (CBP) epitopes M17, M31 and M55 (see
CC AAW31552-54) that confer protection against S. aureus infection. CBP
CC protein and antigenic epitopes are contemplated for use in the treatment
CC of pathological infections, especially to prevent bacterial adhesion to
CC collagen. The epitopes are also contemplated for use in the preparation
CC of vaccines and as carrier proteins in vaccine formulations, as well as
CC in the formulation of compositions for the prevention of S. aureus
CC infection. pCF33 and pQD (see AAW31556) were used to raise anti-MSCRAMM
CC polyclonal antibodies used in passive immunisation against bovine
CC mastitis. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-
CC AUG-2003 to correct OS field.)
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1157; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.le-86;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLMPAYIDPENVKKTGNVTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
DB 125 TLMPAYIDPENVKKTGNVTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 184

QY 61 NTRYQTIVNPSGDNVIAPVLTGNLKTNTSNALIDQONTSIKVKYVDNAADLSSEYFVN 120
DB 185 NTRYQTIVNPSGDNVIAPVLTGNLKTNTSNALIDQONTSIKVKYVDNAADLSSEYFVN 244

QY 121 PENFEDVTNSVNIITFPNPNQYKVFETPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 180
DB 245 PENFEDVTNSVNIITFPNPNQYKVFETPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 304

QY 181 YNSNIIWRSMWDNEVAFNNGSGGSGDGIDKVPVPEQDPE 219
DB 305 YNSNIIWRSMWDNEVAFNNGSGGSGDGIDKVPVPEQDPE 343

RESULT 3
AAE29262
ID AAE29262 standard; protein; 520 AA.
XX
XX AAE29262;
XX
XX 27-JAN-2003 (first entry)
XX
XX Staphylococcus aureus Clf40 protein.
XX
XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.
XX
XX Staphylococcus aureus.
XX
XX WO200272600-A2.
XX
XX 19-SEP-2002.
XX
XX 28-JAN-2002; 2002WO-US002296.
XX
XX 26-JAN-2001; 2001US-0264072P.
PR 12-MAR-2001; 2001US-0274611P.
PR 18-JUN-2001; 2001US-0298413P.
PR 30-JUL-2001; 2001US-0308116P.
XX
XX (INHI-) INHIBITEX INC.
XX
XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
XX WPI; 2002-759834/82.
XX N-PSDB; AAD46861.
XX
XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for
PT treating or preventing Staphylococcus aureus infection e.g. wound
PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in

PT a human or animal.
 XX Claim 9; Page 67-69; 80pp; English.
 XX The invention relates to monoclonal antibody which binds the clumping
 CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
 CC monoclonal antibody is useful for treating or preventing S. aureus
 CC infection in a human or animal, and for inhibiting the binding of
 CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
 CC S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3
 CC protein is useful for inducing an immunological response in a human or
 CC animal. These staphylococcal infections include wound infections, sepsis,
 CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
 CC present sequence is Staphylococcus aureus Clf40 protein
 XX
 SQ Sequence 520 AA;

Query Match 100.0%; Score 1157; DB 5; Length 520;
 Best Local Similarity 100.0%; Pred. No. 1.9e-86;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDIDKTN 60
 DB 293 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDIDKTN 352
 QY 61 NTRQTIYVNPDSGDNVIAPIVLTGNLKPTDSNALIDQNTSIKVKVDNAADLSESYFVN 120
 DB 353 NTRQTIYVNPDSGDNVIAPIVLTGNLKPTDSNALIDQNTSIKVKVDNAADLSESYFVN 412
 QY 121 PENFEDVTNSVNIITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 180
 DB 413 PENFEDVTNSVNIITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 472
 QY 181 YNSNIIWRSMWNEVAFNNGSGGSGGIDKPKVVPQEPDE 219
 DB 473 YNSNIIWRSMWNEVAFNNGSGGSGGIDKPKVVPQEPDE 511

RESULT 4
 ABM72221
 ID ABM72221 standard; protein; 927 AA.
 AC
 XX ABM72221;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Staphylococcus aureus protein #1461.
 XX
 KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 KW enzymatic assay; antibiotic target.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200294868-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 27-MAR-2002; 2002WO-IB002637.
 XX
 PR 27-MAR-2001; 2001GB-00007661.
 XX
 XX (CHIR-) CHIRON SPA.
 XX
 XX Masignani V, Mora M, Scarselli M;
 PI
 XX WPI: 2003-120786/11.
 DR N-PSDB; ACF73781.
 XX
 XX New Staphylococcus aureus protein, useful as a vaccine for treating or
 PT preventing Staphylococcal infection, specifically an infection caused by
 PT S. aureus, e.g. sepsis.
 XX
 PS Claim 1; SEQ ID NO 2922; 49pp; English.

XX The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus proteins of the invention
 XX
 SQ Sequence 927 AA;

Query Match 100.0%; Score 1157; DB 6; Length 927;
 Best Local Similarity 100.0%; Pred. No. 4e-86;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDIDKTN 60
 DB 332 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDIDKTN 391
 QY 61 NTRQTIYVNPDSGDNVIAPIVLTGNLKPTDSNALIDQNTSIKVKVDNAADLSESYFVN 120
 DB 392 NTRQTIYVNPDSGDNVIAPIVLTGNLKPTDSNALIDQNTSIKVKVDNAADLSESYFVN 451
 QY 121 PENFEDVTNSVNIITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 180
 DB 452 PENFEDVTNSVNIITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 511
 QY 181 YNSNIIWRSMWNEVAFNNGSGGSGGIDKPKVVPQEPDE 219
 DB 512 YNSNIIWRSMWNEVAFNNGSGGSGGIDKPKVVPQEPDE 550

RESULT 5
 AAY58435
 ID AAY58435 standard; protein; 933 AA.
 XX
 AC AAY58435;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Staphylococcus aureus fibrinogen binding ClfA protein.
 XX
 KW ClfA; fibrinogen binding protein; bacterial colonisation;
 KW indwelling medical device; staphylococcal infection.
 XX
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..39
 FT /note= "Signal peptide"
 FT Region 40..559
 FT /note= "Region A"
 FT Region 332..550
 FT /note= "Fibrinogen-binding region"
 FT Region 560..867
 FT /note= "Region R"
 FT Region 896..900
 FT /note= "Gram positive wall-associated consensus motif"

US6008341-A.
 XX
 XX 28-DEC-1999.
 XX
 XX 22-AUG-1994; 94US-00293728.
 XX
 XX 22-AUG-1994; 94US-00293728.
 XX
 XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
 XX
 XX Foster TJ, McDevitt DL;

copy print

PF 19-OCT-1999; 99US-00421868.
 XX XX
 PR 22-AUG-1994; 94US-00293728.
 XX XX
 FA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
 XX XX
 PI Foster TJ, Mcdevitt DL;
 XX XX
 XX WPI: 2001-181608/18.
 DR N-PSDB; AAF58593.
 XX XX
 XX Novel Staphylococcus aureus fibrinogen binding protein useful as vaccine
 PT to prevent infection, promote wound healing, block adherence to
 PT indwelling medical devices and for diagnosing staphylococcus aureus
 PT infection.
 XX XX
 XX Claim 5; Fig 2; 30pp; English.
 XX XX
 CC The present sequence is a novel Staphylococcus aureus fibrinogen binding
 CC protein. It is useful as a vaccine to protect against human and animal
 CC infections caused by S.aureus, such as against mastitis, to block
 CC S.aureus from colonising and infecting a wound, to block adherence of
 CC S.aureus to indwelling medical devices such as catheters, replacement
 CC heart valves and cardiac assist devices. The protein can be used to
 CC diagnose bacterial infections. Polyclonal and monoclonal antibodies
 CC raised against the fibrinogen binding protein are useful for passively
 CC immunising against infections caused by S.aureus, to prevent infection of
 CC a wound and to diagnose bacterial infections
 XX XX
 SQ Sequence 933 AA;
 Query Match 100.0%; Score 1157; DB 4; Length 933;
 Best Local Similarity 100.0%; Pred. No. 4e-86;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTMPAYIDPENKKTGNVTLATGISTANKTVLDVYKGYKFNLSIKGFIQIDKTN 60
 DB 332 TLTMPAYIDPENKKTGNVTLATGISTANKTVLDVYKGYKFNLSIKGFIQIDKTN 391
 QY 61 NYRTQTIYVPSGDNVIAFLVLTGNLKPNTDSNALIDQQNTSIKVKVDNAADLSSEYFVN 120
 DB 392 NYRTQTIYVPSGDNVIAFLVLTGNLKPNTDSNALIDQQNTSIKVKVDNAADLSSEYFVN 451
 QY 121 PENFEDVTNSVNIITFPNPQYKVEFNTDDQITTPYIVVNGHIDPNSKGDIALRSTLYG 180
 DB 452 PENFEDVTNSVNIITFPNPQYKVEFNTDDQITTPYIVVNGHIDPNSKGDIALRSTLYG 511
 QY 181 YNSNIIRWSMDNEVAFNNGSGGDIKPVVPEQDE 219
 DB 512 YNSNIIRWSMDNEVAFNNGSGGDIKPVVPEQDE 550
 RESULT 7
 ABJ18947
 ID ABJ18947 standard; protein; 933 AA.
 XX XX
 AC ABJ18947;
 XX XX
 DT 06-MAR-2003 (first entry)
 XX XX
 DE Pathogen specific antigen related staphylococcal protein SEQ ID No 93.
 XX XX
 KW Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
 KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
 KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
 KW autoimmune disease; HIV; hepatitis.
 XX XX
 OS Staphylococcus sp.
 XX XX
 PN WO200259148-A2.
 XX XX
 PD 01-AUG-2002.
 XX XX

XX WPI: 2000-096389/08.
 DR N-PSDB; AA255832.
 XX XX
 XX Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus
 PT aureus, useful for treatment or prevention of infections.
 XX XX
 XX Disclosure; Fig 2A-1-4; 35pp; English.
 PS PS
 XX XX
 CC This sequence represents the Staphylococcus aureus fibrinogen-binding
 CC protein, ClfA. ClfA is an important receptor involved in S. aureus
 CC colonisation of indwelling medical devices (e.g., catheters, artificial
 CC heart valves). Shortly after implantation, the surfaces of medical
 CC devices become coated with host plasma and matrix proteins such as
 CC fibrinogen and fibronectin, and there is considerable evidence to suggest
 CC that bacterial adherence to fibrinogen/fibrin is important in the
 CC initiation of device-related infection. The fibrinogen-binding region of
 CC ClfA is thought to reside between residues 332 and 550 in a region
 CC comprising 154 repeats of the dipeptide Ser-Asp, and the C-terminus
 CC designated A. The protein also contains a repeated region (region R)
 CC contains features present in surface proteins of other Gram positive
 CC bacteria that are responsible for anchoring the protein to the cell wall
 CC and cell membrane. ClfA, or its fragments, may be used to block S. aureus
 CC colonisation of wounds, to prevent adherence of S. aureus to indwelling
 CC medical devices, as vaccines to protect against S. aureus infection
 CC (e.g., mastitis in ruminants), to raise specific antibodies, and for
 CC diagnosis (by agglutination or immunoassay). The specific antibodies are
 CC used for passive immunisation, to block infection of wounds or adhesion
 CC of S. aureus and for diagnosis. Nucleotides encoding ClfA and its
 CC fragments may be used as diagnostic probes
 XX XX
 SQ Sequence 933 AA;
 Query Match 100.0%; Score 1157; DB 3; Length 933;
 Best Local Similarity 100.0%; Pred. No. 4e-86;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTMPAYIDPENKKTGNVTLATGISTANKTVLDVYKGYKFNLSIKGFIQIDKTN 60
 DB 332 TLTMPAYIDPENKKTGNVTLATGISTANKTVLDVYKGYKFNLSIKGFIQIDKTN 391
 QY 61 NYRTQTIYVPSGDNVIAFLVLTGNLKPNTDSNALIDQQNTSIKVKVDNAADLSSEYFVN 120
 DB 392 NYRTQTIYVPSGDNVIAFLVLTGNLKPNTDSNALIDQQNTSIKVKVDNAADLSSEYFVN 451
 QY 121 PENFEDVTNSVNIITFPNPQYKVEFNTDDQITTPYIVVNGHIDPNSKGDIALRSTLYG 180
 DB 452 PENFEDVTNSVNIITFPNPQYKVEFNTDDQITTPYIVVNGHIDPNSKGDIALRSTLYG 511
 QY 181 YNSNIIRWSMDNEVAFNNGSGGDIKPVVPEQDE 219
 DB 512 YNSNIIRWSMDNEVAFNNGSGGDIKPVVPEQDE 550
 RESULT 6
 AAB69508
 ID AAB69508 standard; protein; 933 AA.
 XX XX
 AC AAB69508;
 XX XX
 DT 23-APR-2001 (first entry)
 XX XX
 DE Staphylococcus aureus ClfA protein.
 XX XX
 KW Staphylococcus aureus; clfA; antibiotic; vaccine;
 KW fibrinogen binding protein; bacterial infection; mastitis.
 XX XX
 OS Staphylococcus aureus.
 XX XX
 PN US6177084-B1.
 XX XX
 PD 23-JAN-2001.
 XX XX

PF 21-JAN-2002; 2002WO-EP000546.
XX
PR 26-JAN-2001; 2001AT-00000130.
XX
PA (CISTEM) CISTEM BIOTECHNOLOGIES GMBH.
XX
PI Meinke A, Nagy E, Von Ahlsen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;
XX Tempelmaier B;
XX WPI; 2003-075410/07.
XX
XX Identifying, isolating and producing hyperimmune serum-reactive antigens
PT from a pathogen, for preparing vaccine or medicament for treating or
PT preventing e.g. staphylococcal infections, comprises providing antibody
PT preparation.
XX
XX Example 7; Page 160; 252pp; English.
PS
XX
XX The invention relates to a novel method for identifying, isolating and
CC producing hyperimmune serum-reactive antigens from a pathogen, tumour,
CC allergen, a tissue or host prone to auto-immunity, where the antigens are
CC used in a vaccine, comprises providing antibody preparation from a plasma
CC pool of a type of animal, or individual sera with antibodies against the
CC specific pathogen, tumour, allergen, tissue or host prone to auto-
CC immunity. The hyperimmune serum-reactive antigens comprising any of the
CC 62 sequences of 53-2261 amino acids fully defined in the specification,
CC or their hyperimmune fragments are useful for the manufacture of a
CC pharmaceutical preparation, particularly a vaccine against staphylococcal
CC infections or colonisation against *S. aureus* or *S. epidermidis*. The
CC preparation of antibodies is useful for the manufacture of a medicament
CC for treating or preventing staphylococcal infections or colonisation
CC against *S. aureus* or *S. epidermidis*. The antibody preparations may also
CC be used for diagnostic and imaging purposes. Other conditions that can be
CC treated include cancer, autoimmune diseases or infections caused by viral
CC (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
CC sequence represents a staphylococcal protein relating to the method for
CC identifying and producing pathogen specific antigens of the invention
XX
SQ Sequence 933 AA;

Query Match 100.0%; Score 1157; DB 6; Length 933;
Best Local Similarity 100.0%; Pred. No. 4e-86;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVTATIGSTTANKTVLDYKGYFYNLSIKGTIDQIDKTN 60
Db |||||
332 TLTPAYIDPENVKKTGNVTATIGSTTANKTVLDYKGYFYNLSIKGTIDQIDKTN 391

QY 61 NTYRQIYVNPNGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFVN 120
Db |||||
392 NTYRQIYVNPNGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFVN 451

QY 121 PENFEDVTSNLTFFPNQYKVEFNTPDQITTPYIVVNGHIDNSKGDALRSTLYG 180
Db |||||
452 PENFEDVTSNLTFFPNQYKVEFNTPDQITTPYIVVNGHIDNSKGDALRSTLYG 511

QY 181 YNSNIIWRSMDSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
Db |||||
512 YNSNIIWRSMDSWDNEVAFNNGSGSGDGIDKPVVPEQDE 550

RESULT 8
AAW89801
ID AAW89801 standard; protein; 936 AA.
XX
XX AAW89801;
XX
DT 16-MAR-1999 (first entry)
XX
DE Staphylococcus aureus protein SEQ ID #5249.
XX
KW Computer readable medium; vaccine; *S. aureus* infection; immunodetection;

KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
XX toxic shock syndrome.
XX Staphylococcus aureus.
XX EP786519-A2.
XX 30-JUL-1997.
XX
XX 07-JAN-1997; 97EP-00100117.
XX
XX 05-JAN-1996; 96US-0009861P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
XX WPI; 1997-374922/35
XX
XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -
PT stored on computer readable medium and used in the production of anti-
PT *S. aureus* vaccines.
XX
XX Claim 23; Page 3255-3258; 3271pp; English.
XX
XX This sequence represents a *Staphylococcus aureus* protein sequence of the
CC invention. The DNA sequences encoding the *S. aureus* proteins are recorded
CC on a computer readable medium, preferably selected from a floppy or hard
CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.
CC Homology searches using the *S. aureus* DNA sequences allows putative
CC functions to be assigned so that protein-encoding or regulatory regions
CC of commercial, therapeutic or industrial importance can be obtained.
CC Specifically, sequences which are likely to encode antigens have been
CC identified and these polypeptides can be used in a vaccine composition
CC against *S. aureus* infection. The polypeptides can also be used in a kit
CC for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated
CC in numerous human diseases, including cellulitis, eyelid infections, food
CC poisoning, osteomyelitis, skin and surgical wound infections, scalded
CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the
CC DNA sequences can be used for recombinant production of the polypeptides.
CC The new DNA sequences (and their fragments) are useful as primers or
CC probes for isolating homologues of any of the 5191 *S. aureus* DNA sequences
XX contained on the computer readable medium
SQ Sequence 936 AA;

Query Match 100.0%; Score 1157; DB 2; Length 936;
Best Local Similarity 100.0%; Pred. No. 4e-86;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVTATIGSTTANKTVLDYKGYFYNLSIKGTIDQIDKTN 60
Db |||||
341 TLTPAYIDPENVKKTGNVTATIGSTTANKTVLDYKGYFYNLSIKGTIDQIDKTN 400

QY 61 NTYRQIYVNPNGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFVN 120
Db |||||
401 NTYRQIYVNPNGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFVN 460

QY 121 PENFEDVTSNLTFFPNQYKVEFNTPDQITTPYIVVNGHIDNSKGDALRSTLYG 180
Db |||||
461 PENFEDVTSNLTFFPNQYKVEFNTPDQITTPYIVVNGHIDNSKGDALRSTLYG 520

QY 181 YNSNIIWRSMDSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
Db |||||
521 YNSNIIWRSMDSWDNEVAFNNGSGSGDGIDKPVVPEQDE 559

RESULT 9
ABU16402
ID ABU16402 standard; protein; 935 AA.
XX
XX ABU16402;

QY 2 LTWPAYIDPENVKKTGNVTLATGIGSTANKTVLADYKEYKGFYNLSIKGTIDQIDKINN 61
Db 333 ITWPAYIDPENVTKTGNVTLTGTGINTASKTVLIDYKEYGQFHNLSIKGTIDQIDKINN 392
QY 62 TYRQTIYVNPSSGDNVIAFVLTLGNLKPNTDSNALIDQOQNTSIKVKYKVDNAADLSSEYFVNP 121
Db 393 TYRQTIYVNPSSGDNVLPALTGNLIFNTKSNALIDAKNTDIKVVYVDNANDLSSEYVNP 452
QY 122 ENFEDVTNSVNIITFPNPNQYKVEFTEPDQITTTTPIYVWNGHIDPNKGDALBSTLYGY 181
Db 453 SDFEDVTNQVRIISFPNANQYKVEFTEPDQITTTTPIYVWNGHIDFASGTALRSTFYGY 512
QY 182 NSNIIRWSMSWDNEVAFNNGSGDGDIDKPVVPQDPE 219
Db 513 DSNFIWRMSWDNEVAFNNGSGDGDIDKPVVPQDPE 550
RESULT 10
AAU33975
ID AAU33975 standard; protein; 1021 AA.
XX
AC AAU33975;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #251.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207272P.
PR 23-OCT-2000; 2000US-0243578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS51834.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 5471; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic

XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #1929.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA20272.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 44326; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC a gene on which the test compound that inhibits proliferation of the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC identifying a gene required for cellular proliferation or the biological
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 935 AA;
Query Match 85.7%; Score 991; DB 6; Length 935;
Best Local Similarity 84.9%; Pred. No. 1.7e-72;
Matches 185; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1021 AA;

Query Match	85.7%;	Score 991;	DB 4;	Length 1021;
Best Local Similarity	84.9%;	Pred. No. 1.9e-72;		
Matches 185;	Conservative 13;	Mismatches 20;	Indels 0;	Gaps 0
Qy	2	LTMPAYIDPENVKVTGNVLTATIGSTTANKTVLVDVDEYKGYKGFYNLSIKGHIQIDQDKTNN	61	
Db	333	ITMTPAYIDPENVTGNTVLTGTTGVTWASKTVLVDIYKGYQGFHLSIKGHIQIDQDKTNN	392	
Qy	62	TYRQTIYVNSGDNVIAPVLTGNLKPNTDSNALIDQOQNTSIKVKYVDNAALDSYSFYVNP	121	
Db	393	TYRQTIYVNSGDNVLPALPTGNLIPNTKSNALIDAKNTDIKVKYVDNANDLSESYVNP	452	
Qy	122	ENFEDVTNSVNIITPPENQYKVEENTPDDQITTPYIVVNGHIDPNSKGDLALRSTLYGY	181	
Db	453	SDFEDVTNQVRSFPNANQYKVEPTDDQITTPYIVVNGHIDPASTGDUALRSTFYGY	512	
Qy	182	NSNIIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDPE	219	
Db	513	DSNFIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDPE	550	

RESULT 11	
AAU36951	
ID	AAU36951 standard; protein; 1021 AA.
XX	
AC	AAU36951;
XX	
DT	14-FEB-2002 (first entry)
XX	
DE	Staphylococcus aureus cellular proliferation protein #1121.
XX	
KW	Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW	antibacterial; drug design.
XX	
OS	Staphylococcus aureus.

~~(S)~~

WO200170955-A2.
27-SEP-2001.
21-MAR-2001; 2001WO-US009180.

21-MAR-2000; 2000US-0191078P.
23-MAY-2000; 2000US-0206848P.
26-MAY-2000; 2000US-0207727P.
23-OCT-2000; 2000US-0242578P.
27-NOV-2000; 2000US-0253625P.
22-DEC-2000; 2000US-0257931P.
16-FEB-2001; 2001US-0269308P.

(ELIT-) ELITRA PHARM INC.
Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
yamamoto RT, Xu HH;
WPI; 2001-611495/70.
N-PADB; AAS54810.

X
X
PN
XX
PPD
XX
PF
XX
PER

New polynucleotides for the identification and development of antibiotics, comprise sequences of antisense nucleic acids.

Example 3; SEQ ID NO 12544; 511pp; English.

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes

themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published pct sequences](http://wipo.int/pub/published/pct/sequences)

Query Match	85.7%	Score 991;	DB 4;	Length 1021;
Best Local Similarity	84.9%;	Pred. No. 1.9e-72;		
Matches 185;	Conservative 13;	Mismatches 20;	Indels 0;	Gaps 0
QY	2	LTPPAYIDPENVKKTGNVTLATGIGSTANKTVLVDIEKYGKFVNLISIKGHIQIDIKTKNN	61	
Db	333	ITWPAVIDPENVTGKGNVTLTGIGTASTKVLIDIEKYGQGFNLISIKGTIDQIDIKTKNN	392	
QY	62	TYRQTIYVNPSPGDNVIAPVLITGNLKPNSTDNALIDQONTISIKVKYVDNAADLSSESYFVNP	121	
Db	393	TYRQTIYVNPSPGDNVVLPAITGNLIPNTKSNALIDAKNTDIKVRVDNANDLSSESYVNP	452	
QY	122	ENPEDVTNSVNIETFPNPQYKVFENPTDDQITTPYIVVNGHIDPNSGDLALRSTLYGY	181	
Db	453	SFEDVTNQVRISFPFNANQYKVFPTDDQITTPYIVVNGHIDPASTGDLALRSTFYGY	512	
QY	182	NSNIINRWSWNEVAFNNGSGSGDGIDKPVPPEQDE	219	
Db	513	DSNFIINRWSWNEVAFNNGSGSGDGIDKPVPPEQDE	550	

RESULT 12	
AAU75490	
ID	AAU75490 standard; protein; 496 AA.
XX	
XX	AAU75490;
XX	
XX	
DT	23-APR-2002 (first entry)
XX	
XX	S. aureus antigenic protein associated protein #10.
DE	
XX	
KW	Antigenic protein; vaccine; SEREX; antibacterial; antiinflammatory;
KW	dermatological; antiulcer; tuberculostatic; immunosuppressive;
KW	septicemia; food poisoning; skin disorders; peritonitis; endocarditis;
KW	tuberculosis; blood infection; sepsis; meningitis; pneumonia;
KW	stomach ulcer; gonorrhea; necrotizing fasciitis; impetigo;
KW	Lyme's disease; gastro-enteritis; dysentery; shigellosis.
XX	
OS	Staphylococcus aureus.
XX	
FPN	WO200198499-A1.
XX	
XX	27-DEC-2001.
XX	
XX	20-JUN-2001; 2001WO-GB002685.
XX	
XX	20-JUN-2000; 2000GB-00014907.
XX	
XX	(UYSH-) UNIV SHEFFIELD.
XX	(BIOS-) BIOSYNEXUS INC.
XX	
PPI	Foster S, McDowell P, Brummell K, Clarke S;
XX	
XX	WPI; 2002-106544/14.

PT Identifying antigenic polypeptides expressed by pathogenic organisms
PT e.g., Staphylococcus aureus during infection, by SEREX (serological
PT identification of antigens by recombinant expression cloning) techniques.
XX
XX
PS Disclosure; Page 72-74; 85pp; English.

XX The invention relates to a method for identifying antigenic polypeptides
XX expressed by pathogenic organisms e.g., Staphylococcus aureus during
XX infection, by SEREX (serological identification of antigens by
XX recombinant expression cloning) techniques. The method involves providing
XX a nucleic acid library encoding genes/partial gene sequences of
XX pathogenic organisms, transforming/transfecting the library into host
XX cells, contacting the polypeptides expressed by the genes with autologous
XX antisera derived from an animal infected with, or has been infected with
XX the pathogen and purifying the nucleic acid encoding the polypeptide or
XX partial polypeptide binding to the antisera. Also included are the
XX nucleic acids and polypeptides isolated by the method, vectors and
XX transformed cells expressing them, a vaccine comprising the polypeptide
XX and the production of monoclonal antibodies against the polypeptides. The
XX protein and vaccine are useful for immunising an animal (preferably
XX human) against a pathogenic microbe. The proteins and antibodies are
XX useful for manufacturing a medicament for treating Staphylococcus aureus-
XX associated septicaemia, food poisoning or skin disorders or
XX Staphylococcus epidermidis-associated septicaemia, sepsis, meningitis,
XX endocarditis, tuberculosis, blood infections, sepsis, meningitis,
XX pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, impetigo,
XX Lyme's disease, gastro-enteritis, dysentery and shigellosis. The present
XX sequence is an S. aureus protein sequence associated with the antigenic
XX proteins of the invention. Note: The present sequence is included in the
XX sequence listing but is not mentioned anywhere else in the specification
XX
XX Sequence 496 AA;

Query Match 74.6%; Score 863; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.5e-62;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVLTATGISTTANKTGLVLDYKGYKFNLSIKGTIDQDKTN 60
Db 332 TLTPAYIDPENVKKTGNVLTATGISTTANKTGLVLDYKGYKFNLSIKGTIDQDKTN 391
QY 61 NTVRQIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVN 120
Db 392 NTVRQIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVN 451
QY 121 PENFEDVTNSVNIFFPNPNQYKVEFNTPDQITTPYIVVNGHID 165
Db 452 PENFEDVTNSVNIFFPNPNQYKVEFNTPDQITTPYIVVNGHID 496

RESULT 13
ADA89664
ID ADA89664 standard; protein; 496 AA.
XX
XX ADA89664;
XX
XX 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus antigenic protein #203.
XX
XX antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
XX antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
XX antiulcer; immunostimulant; ophthalmological; pathogenic microbe;
XX bacteriaemia; septic shock; organ infection; skin infection;
XX bacterial basal colonisation; bacterial eye infection; septicaemia;
XX tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
XX sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
XX necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
XX gastro-enteritis; dysentery; shigellosis; skin disorder.
XX
XX Staphylococcus aureus.
XX
XX WO2003011899-A2.
XX
XX

XX 13-FEB-2003.
XX
XX 02-AUG-2002; 2002WO-GB003606.
XX
XX 02-AUG-2001; 2001GB-00018825.
XX 09-JAN-2002; 2002GB-00000349.
XX (UYSH-) UNIV SHEFFIELD.
XX (BIOS-) BIOSYNEXUS INC.
XX Foster S, Mord J, Clarke S, Medowell P, Brummel K;
XX WPI; 2003-256434/25.
XX
XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
XX useful as a vaccine for immunizing humans against e.g. bacteraemia, septic
XX shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
XX impetigo.
XX
XX Claim 4; Page 161; 189pp; English.

XX The present invention describes an antigenic protein or its part, which
XX is for use as a vaccine. The antigenic protein is encoded by an isolated
XX DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
XX sequences (designated dnaSA and dna SE, respectively; and which encodes a
XX protein expressed by a pathogenic organism. Also described: (1) a vaccine
XX composition comprising at least one antigenic protein; (2) a method of
XX immunising an animal against a disease or condition caused by a
XX pathogenic microbe by administering the antigenic protein or the vaccine;
XX (3) an antibody or its binding part obtainable by the method above; (4)
XX preparing a hybridoma cell line producing monoclonal antibodies; (5) a
XX hybridoma cell line produced by the method of (4); and (6) identifying
XX opsonic antigens expressed by a pathogenic microbe. The antigenic
XX proteins have antibacterial, neuroprotective, immunosuppressive,
XX antiinflammatory, antiulcer, immunostimulant and ophthalmological
XX activities, and can be used in vaccines. The antigenic proteins or
XX vaccines can be used for immunising an animal (specifically a human)
XX against a disease or condition caused by a pathogenic microbe, e.g.
XX bacteraemia, septic shock, organ infection, skin infection, bacterial
XX basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
XX bacteria-associated food poisoning, blood infections, peritonitis,
XX endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
XX strep throat, streptococcal-associated toxic shock, necrotising
XX fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
XX dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
XX skin disorders, S. epidermidis-associated septicaemia, peritonitis or
XX endocarditis. The present sequence represents a S. aureus antigenic
XX protein sequence from the present invention.

XX Sequence 496 AA;
Query Match 74.6%; Score 863; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.5e-62;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVLTATGISTTANKTGLVLDYKGYKFNLSIKGTIDQDKTN 60
Db 332 TLTPAYIDPENVKKTGNVLTATGISTTANKTGLVLDYKGYKFNLSIKGTIDQDKTN 391
QY 61 NTVRQIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVN 120
Db 392 NTVRQIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVN 451
QY 121 PENFEDVTNSVNIFFPNPNQYKVEFNTPDQITTPYIVVNGHID 165
Db 452 PENFEDVTNSVNIFFPNPNQYKVEFNTPDQITTPYIVVNGHID 496
XX
XX RESULT 14
XX ADA89663
XX ID ADA89663 standard; protein; 496 AA.
XX

Query Match 74.6%; Score 863; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.5e-62;

CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published pct sequences

ftp.wipo.int/pub/published_pct_sequences
XX

XX

Sequence 767 AA;

Query Match 22.0%; Score 255; DB 4; Length 767;

Query Match	22.0%;	Score 435;	8;
Best Local Similarity	26.9%;	Pred. No. 3.6e-12;	
Matches	61;	Conservative 54;	Mismatches 82;
Indels	30;	Gaps	8;

Matches	61;	Conservative	54;	Mismatches	82;	Indels	30;	Gaps	8;
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56

273 I:SI:RI:DP:TT:V:OK:NON:VE:KL:GET:V:SK:FI:NI:O:VL:GV:RD:NG:---VT:ANG:ID:TL 328

57 DKTNNTYBOTTVVNPSGDNVAPVLTGNL----KPNTDSNALIDQONTSIKVYKVDNAAD 112

329 NKVDGFESHFAVYKPNNOCSLSVTVTGVTGKNKPGVN-----NPTVKYKHIGSDD 380

113 ISES YFVN PEN --- FED VTS VN IT FPN PN QY KVE FNT PDD Q IT TP YI VV VNG H ID P NSK 169

QY 113 LBSB11747EN 425 140 142 144 146 148 150 152 154 156 158 160 162 164 166 168 170 172 174 176 178 180 182 184 186 188 190 192 194 196 198 200 202 204 206 208 210 212 214 216 218 220 222 224 226 228 230 232 234 236 238 240 242 244 246 248 250 252 254 256 258 260 262 264 266 268 270 272 274 276 278 280 282 284 286 288 290 292 294 296 298 300 302 304 306 308 310 312 314 316 318 320 322 324 326 328 330 332 334 336 338 340 342 344 346 348 350 352 354 356 358 360 362 364 366 368 370 372 374 376 378 380 382 384 386 388 390 392 394 396 398 400 402 404 406 408 410 412 414 416 418 420 422 424 426 428 430 432 434 436 438 440 442 444 446 448 450 452 454 456 458 460 462 464 466 468 470 472 474 476 478 480 482 484 486 488 490 492 494 496 498 500 502 504 506 508 510 512 514 516 518 520 522 524 526 528 530 532 534 536 538 540 542 544 546 548 550 552 554 556 558 560 562 564 566 568 570 572 574 576 578 580 582 584 586 588 590 592 594 596 598 600 602 604 606 608 610 612 614 616 618 620 622 624 626 628 630 632 634 636 638 640 642 644 646 648 650 652 654 656 658 660 662 664 666 668 670 672 674 676 678 680 682 684 686 688 690 692 694 696 698 700 702 704 706 708 710 712 714 716 718 720 722 724 726 728 730 732 734 736 738 740 742 744 746 748 750 752 754 756 758 760 762 764 766 768 770 772 774 776 778 780 782 784 786 788 790 792 794 796 798 800 802 804 806 808 810 812 814 816 818 820 822 824 826 828 830 832 834 836 838 840 842 844 846 848 850 852 854 856 858 860 862 864 866 868 870 872 874 876 878 880 882 884 886 888 890 892 894 896 898 900 902 904 906 908 910 912 914 916 918 920 922 924 926 928 930 932 934 936 938 940 942 944 946 948 950 952 954 956 958 960 962 964 966 968 970 972 974 976 978 980 982 984 986 988 990 992 994 996 998 1000 1002 1004 1006 1008 1010 1012 1014 1016 1018 1020 1022 1024 1026 1028 1030 1032 1034 1036 1038 1040 1042 1044 1046 1048 1050 1052 1054 1056 1058 1060 1062 1064 1066 1068 1070 1072 1074 1076 1078 1080 1082 1084 1086 1088 1090 1092 1094 1096 1098 1100 1102 1104 1106 1108 1110 1112 1114 1116 1118 1120 1122 1124 1126 1128 1130 1132 1134 1136 1138 1140 1142 1144 1146 1148 1150 1152 1154 1156 1158 1160 1162 1164 1166 1168 1170 1172 1174 1176 1178 1180 1182 1184 1186 1188 1190 1192 1194 1196 1198 1200 1202 1204 1206 1208 1210 1212 1214 1216 1218 1220 1222 1224 1226 1228 1230 1232 1234 1236 1238 1240 1242 1244 1246 1248 1250 1252 1254 1256 1258 1260 1262 1264 1266 1268 1270 1272 1274 1276 1278 1280 1282 1284 1286 1288 1290 1292 1294 1296 1298 1300 1302 1304 1306 1308 1310 1312 1314 1316 1318 1320 1322 1324 1326 1328 1330 1332 1334 1336 1338 1340 1342 1344 1346 1348 1350 1352 1354 1356 1358 1360 1362 1364 1366 1368 1370 1372 1374 1376 1378 1380 1382 1384 1386 1388 1390 1392 1394 1396 1398 1400 1402 1404 1406 1408 1410 1412 1414 1416 1418 1420 1422 1424 1426 1428 1430 1432 1434 1436 1438 1440 1442 1444 1446 1448 1450 1452 1454 1456 1458 1460 1462 1464 1466 1468 1470 1472 1474 1476 1478 1480 1482 1484 1486 1488 1490 1492 1494 1496 1498 1500 1502 1504 1506 1508 1510 1512 1514 1516 1518 1520 1522 1524 1526 1528 1530 1532 1534 1536 1538 1540 1542 1544 1546 1548 1550 1552 1554 1556 1558 1560 1562 1564 1566 1568 1570 1572 1574 1576 1578 1580 1582 1584 1586 1588 1590 1592 1594 1596 1598 1600 1602 1604 1606 1608 1610 1612 1614 1616 1618 1620 1622 1624 1626 1628 1630 1632 1634 1636 1638 1640 1642 1644 1646 1648 1650 1652 1654 1656 1658 1660 1662 1664 1666 1668 1670 1672 1674 1676 1678 1680 1682 1684 1686 1688 1690 1692 1694 1696 1698 1700 1702 1704 1706 1708 1710 1712 1714 1716 1718 1720 1722 1724 1726 1728 1730 1732 1734 1736 1738 1740 1742 1744 1746 1748 1750 1752 1754 1756 1758 1760 1762 1764 1766 1768 1770 1772 1774 1776 1778 1780 1782 1784 1786 1788 1790 1792 1794 1796 1798 1800 1802 1804 1806 1808 1810 1812 1814 1816 1818 1820 1822 1824 1826 1828 1830 1832 1834 1836 1838 1840 1842 1844 1846 1848 1850 1852 1854 1856 1858 1860 1862 1864 1866 1868 1870 1872 1874 1876 1878 1880 1882 1884 1886 1888 1890 1892 1894 1896 1898 1900 1902 1904 1906 1908 1910 1912 1914 1916 1918 1920 1922 1924 1926 1928 1930 1932 1934 1936 1938 1940

170 GDIALRSTLYGYNSTLIWRSMWDNEVAFNNGSGDGIDK---PVV 213

[illegible]

Search completed: November 3, 2004, 19:03:02
Job time : 41.4791 secs

Job time : 41.4791 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein ~ protein search, using sw model

Run on: November 3, 2004, 18:54:26 ; Search time 10.7401 Seconds
(without alignments)
1352.277 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_550

Perfect score: 1157

Sequence: 1 TLTMPAYIDPENVKKTGNVT.....NGSGSGDGIDKPVVPEQFDE 219

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1157	100.0	345	3	US-08-856-253-7
2	1157	100.0	933	3	US-08-293-728-2
3	1157	100.0	933	3	US-09-421-868-2
4	1157	100.0	936	4	US-08-956-171E-5249
5	1157	100.0	936	4	US-08-781-986A-5249
6	244.5	21.1	930	3	US-09-134-001C-5314
7	244.5	21.1	930	4	US-09-386-962C-10
8	241.5	20.9	582	4	US-09-147-405B-13
9	241.5	20.9	593	4	US-09-147-405B-11
10	241.5	20.9	1092	4	US-09-147-405B-15
11	208	18.0	918	4	US-09-200-650E-7
12	207.5	17.9	1166	4	US-09-200-650E-1
13	180.5	15.6	1315	4	US-09-200-650E-5
14	175.5	15.2	1027	4	US-08-956-171E-5254
15	175.5	15.2	1027	4	US-08-781-986A-5251
16	136	11.8	559	4	US-08-956-171E-5251
17	136	11.8	559	4	US-08-781-986A-5251
18	128.5	11.1	251	4	US-08-956-171E-5252
19	128.5	11.1	251	4	US-08-781-986A-5252
20	123	10.6	1112	2	US-08-714-402-2
21	123	10.6	1161	3	US-09-327-536-2
22	117.5	10.2	930	4	US-09-200-650E-3
23	112	9.7	1155	4	US-09-710-279-1780
24	112	9.7	1742	4	US-09-386-962C-4
25	112	9.7	1742	4	US-09-386-959-4
26	104.5	9.0	1216	4	US-09-134-000C-5130
27	104.5	9.0	2780	4	US-10-220-587-2

28	104	9.0	1132	4	US-09-248-796A-15026	Sequence 15026, A
29	103	8.9	405	4	US-09-328-352-5300	Sequence 5300, Ap
30	102.5	8.9	559	4	US-09-710-279-1376	Sequence 1376, Ap
31	100.5	8.7	1454	4	US-09-673-896-2	Sequence 2, Appli
32	98	8.5	1457	4	US-09-673-896-4	Sequence 4, Appli
33	97.5	8.4	186	2	US-08-732-791-1	Sequence 1, Appli
34	97.5	8.4	186	3	US-09-205-049-1	Sequence 1, Appli
35	97.5	8.4	1180	4	US-09-206-942-65	Sequence 65, Appl
36	97.5	8.4	1188	4	US-09-206-942-63	Sequence 63, Appl
37	94.5	8.2	992	4	US-09-206-942-61	Sequence 61, Appl
38	94.5	8.2	998	4	US-09-206-942-59	Sequence 59, Appl
39	94.5	8.2	1359	4	US-09-134-000C-6374	Sequence 6374, Ap
40	94	8.1	1095	4	US-09-206-942-45	Sequence 45, Appl
41	94	8.1	1101	4	US-09-206-942-43	Sequence 43, Appl
42	93.5	8.1	526	2	US-08-853-659A-40	Sequence 40, Appl
43	93.5	8.1	2777	4	US-10-220-587-4	Sequence 4, Appli
44	93	8.0	668	4	US-09-277-431A-2	Sequence 2, Appli
45	92.5	8.0	1228	4	US-09-463-402-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210

COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,253
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,678
FILING DATE: 16-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: TAMK:193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-856-253-7

Query Match 100.0%; Score 1157; DB 3; Length 345;
Best Local Similarity 100.0%; Pred.No. 3.4e-105;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match 100.0%; Score 1157; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. NO. 1.5e-104; Indels 0; Gaps 0;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 125 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 184

QY 61 NTRYQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVN 120
DB 185 NTRYQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVN 244

QY 121 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGLALRSTLYG 180
DB 245 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGLALRSTLYG 304

QY 181 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
DB 305 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 343

RESULT 2
US-08-293-728-2
Sequence 2, Application US/08293728D
Patent No. 6008341
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
APPLICANT: McDevitt, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/08/293,728D
CURRENT FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match 100.0%; Score 1157; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. NO. 1.5e-104; Indels 0; Gaps 0;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 391

QY 61 NTRYQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVN 120
DB 392 NTRYQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVN 451

QY 121 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGLALRSTLYG 180
DB 452 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGLALRSTLYG 511

QY 181 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
DB 512 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 550

RESULT 3
US-09-421-868-2
Sequence 2, Application US/09421868
Patent No. 6177084
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
APPLICANT: McDevitt, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/09/421,868
CURRENT FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 08/293,728
PRIOR FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match 100.0%; Score 1157; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. NO. 1.5e-104; Indels 0; Gaps 0;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 125 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 184

QY 61 NTRYQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVN 120
DB 185 NTRYQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVN 244

QY 121 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGLALRSTLYG 180
DB 245 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGLALRSTLYG 304

QY 181 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
DB 305 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 343

RESULT 2
US-08-293-728-2
Sequence 2, Application US/08293728D
Patent No. 6008341
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
APPLICANT: McDevitt, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/08/293,728D
CURRENT FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match 100.0%; Score 1157; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. NO. 1.5e-104; Indels 0; Gaps 0;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 391

QY 61 NTRYQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVN 120
DB 392 NTRYQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVN 451

QY 121 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGLALRSTLYG 180
DB 452 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGLALRSTLYG 511

QY 181 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
DB 512 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 550

RESULT 3
US-09-421-868-2
Sequence 2, Application US/09421868
Patent No. 6177084
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
APPLICANT: McDevitt, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/09/421,868
CURRENT FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 08/293,728
PRIOR FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0



```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5249:
US-08-956-171E-5249

Query Match          100.0%; Score 1157; DB 4; Length 936;
Best Local Similarity 100.0%; Pred. No. 1.5e-104;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
Db 341 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 400

QY 61 NTRYQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQQTSTSIKVKVDNAADLSSEYFVN 120
Db 401 NTRYQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQQTSTSIKVKVDNAADLSSEYFVN 460

QY 121 PENFEDVTVNSNITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGLALRSTLYG 180
Db 461 PENFEDVTVNSNITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGLALRSTLYG 520

QY 181 YNSNIIWRSMWDEVAFNNGSGSGDGIDKPVVPEQDPE 219
Db 521 YNSNIIWRSMWDEVAFNNGSGSGDGIDKPVVPEQDPE 559

RESULT 5
US-08-781-986A-5249
; Sequence 5249, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB2499P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5249:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-781-986A-5249

Query Match          100.0%; Score 1157; DB 4; Length 936;
Best Local Similarity 100.0%; Pred. No. 1.5e-104;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
Db 341 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 400

QY 61 NTRYQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQQTSTSIKVKVDNAADLSSEYFVN 120
Db 401 NTRYQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQQTSTSIKVKVDNAADLSSEYFVN 460

QY 121 PENFEDVTVNSNITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGLALRSTLYG 180
Db 461 PENFEDVTVNSNITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGLALRSTLYG 520

QY 181 YNSNIIWRSMWDEVAFNNGSGSGDGIDKPVVPEQDPE 219
Db 521 YNSNIIWRSMWDEVAFNNGSGSGDGIDKPVVPEQDPE 559

RESULT 6
US-09-134-001C-5314
; Sequence 5314, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5314
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5314

Query Match          21.1%; Score 244.5; DB 3; Length 930;
Best Local Similarity 31.9%; Pred. No. 4.3e-15;
Matches 72; Conservative 36; Mismatches 93; Indels 25; Gaps 10;

QY 2 LTWPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDK 58
Db 386 LKUTSYIDKSKVNNNTKLDVEYKTALSS--VNKTITVEYQKPNENRTANLQSMFTNIDT 443

QY 59 TNNTRYQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQQTSTSIKVKVDNAADLSSEYF 118
Db 444 KNHTVEQTIYINPL--RYSAKETNVAISGNGDEGSGTIIDSTIIKVKYKVDNQNLDPDSNR 501

QY 119 V-NPENFEDVTVNSNITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGD----- 171
Db 502 IYDYSEYEDVTNDVYAQLGNNDVNINFG---NIDSPYIIKVISKYDPN-KDYTTIIQQ 556

QY 172 -LALRSTLYGNSNIIWRSMWDEVAFNNGSGSGDGIDKPVVPEQ 216
Db 557 TVTMQTTINEYTG--FRTASYDNTIAFTSSGQGGQ-DLP--PEK 597

RESULT 7
US-09-386-962C-10
; Sequence 10, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; CURRENT FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
```

```
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-386-962C-10

Query Match      21.1%; Score 244.5; DB 4; Length 930;
Best Local Similarity 31.9%; Pred. No. 4.3e-15;
Matches 72; Conservative 36; Mismatches 93; Indels 25; Gaps 10;

QY 2 LTMPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDK 58
Db 386 LKLTSYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQRPENRTANLQSMFTNIDT 443
QY 59 TNNYTRQTIYVNPBGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSSEYF 118
Db 444 KNHIVEQTIYINPL--RYSAKETNVNISGNGDEGSTIIDDDSTIIKVKYKVDNQNLPSNR 501
QY 119 V-NPENFEDVTNSVNITFPNPQYKVFENTDDQITTPYIVVNGHIDPNSKGD----- 171
Db 502 IVDYSEYEDVTNDYDQAQGNNDVNINFG---NIDSPYIIKVKISKYDPN-KDDYTTIIQQ 556
QY 172 -LALRSTLYGYSNLIWRSMWNEVAFNNGSGDGDIDKPVVPEQ 216
Db 557 TVTMQTITNEYTGE--FRTASYDNTIAFSTSSGQGG--DLP--PEK 597

RESULT 8
US-09-147-405B-13
; Sequence 13, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:
; APPLICANT: Guss, Bengt
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Flock, Jan-Ingmar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; FILE REFERENCE: guss 09/147405
; CURRENT APPLICATION NUMBER: US/09/147,405B
; PRIOR FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-13

Query Match      20.9%; Score 241.5; DB 4; Length 582;
Best Local Similarity 31.4%; Pred. No. 4.2e-15;
Matches 71; Conservative 37; Mismatches 93; Indels 25; Gaps 10;

QY 2 LTMPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDK 58
Db 315 LKLTSYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQRPENRTANLQSMFTNIDT 372
QY 59 TNNYTRQTIYVNPBGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSSEYF 118
Db 373 KNHIVEQTIYINPL--RYSAKETNVNISGNGDEGSTIIDDDSTIIKVKYKVDNQNLPSNR 430
QY 119 V-NPENFEDVTNSVNITFPNPQYKVFENTDDQITTPYIVVNGHIDPNSKGD----- 171
Db 431 IVDYSEYEDVTNDYDQAQGNNDVNINFG---NIDSPYIIKVKISKYDPN-KDDYTTIIQQ 485

; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-386-962C-10

Query Match      21.1%; Score 244.5; DB 4; Length 930;
Best Local Similarity 31.9%; Pred. No. 4.3e-15;
Matches 72; Conservative 36; Mismatches 93; Indels 25; Gaps 10;

QY 2 LTMPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDK 58
Db 386 LKLTSYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQRPENRTANLQSMFTNIDT 443
QY 59 TNNYTRQTIYVNPBGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSSEYF 118
Db 444 KNHIVEQTIYINPL--RYSAKETNVNISGNGDEGSTIIDDDSTIIKVKYKVDNQNLPSNR 501
QY 119 V-NPENFEDVTNSVNITFPNPQYKVFENTDDQITTPYIVVNGHIDPNSKGD----- 171
Db 502 IVDYSEYEDVTNDYDQAQGNNDVNINFG---NIDSPYIIKVKISKYDPN-KDDYTTIIQQ 556
QY 172 -LALRSTLYGYSNLIWRSMWNEVAFNNGSGDGDIDKPVVPEQ 216
Db 557 TVTMQTITNEYTGE--FRTASYDNTIAFSTSSGQGG--DLP--PEK 597

RESULT 8
US-09-147-405B-13
; Sequence 13, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:
; APPLICANT: Guss, Bengt
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Flock, Jan-Ingmar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; FILE REFERENCE: guss 09/147405
; CURRENT APPLICATION NUMBER: US/09/147,405B
; PRIOR FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-13

Query Match      20.9%; Score 241.5; DB 4; Length 582;
Best Local Similarity 31.4%; Pred. No. 4.2e-15;
Matches 71; Conservative 37; Mismatches 93; Indels 25; Gaps 10;

QY 2 LTMPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDK 58
Db 315 LKLTSYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQRPENRTANLQSMFTNIDT 372
QY 59 TNNYTRQTIYVNPBGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSSEYF 118
Db 373 KNHIVEQTIYINPL--RYSAKETNVNISGNGDEGSTIIDDDSTIIKVKYKVDNQNLPSNR 430
QY 119 V-NPENFEDVTNSVNITFPNPQYKVFENTDDQITTPYIVVNGHIDPNSKGD----- 171
Db 431 IVDYSEYEDVTNDYDQAQGNNDVNINFG---NIDSPYIIKVKISKYDPN-KDDYTTIIQQ 485

QY 172 -LALRSTLYGYSNLIWRSMWNEVAFNNGSGDGDIDKPVVPEQ 216
Db 486 TVTMQTITNEYTGE--FRTASYDNTIAFSTSSGQGG--DLP--PEK 526

RESULT 9
US-09-147-405B-11
; Sequence 11, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:
; APPLICANT: Guss, Bengt
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Flock, Jan-Ingmar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; FILE REFERENCE: guss 09/147405
; CURRENT APPLICATION NUMBER: US/09/147,405B
; PRIOR FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-11

Query Match      20.9%; Score 241.5; DB 4; Length 593;
Best Local Similarity 31.4%; Pred. No. 4.3e-15;
Matches 71; Conservative 37; Mismatches 93; Indels 25; Gaps 10;

QY 2 LTMPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDK 58
Db 322 LKLTSYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQRPENRTANLQSMFTNIDT 379
QY 59 TNNYTRQTIYVNPBGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSSEYF 118
Db 380 KNHIVEQTIYINPL--RYSAKETNVNISGNGDEGSTIIDDDSTIIKVKYKVDNQNLPSNR 437
QY 119 V-NPENFEDVTNSVNITFPNPQYKVFENTDDQITTPYIVVNGHIDPNSKGD----- 171
Db 438 IVDYSEYEDVTNDYDQAQGNNDVNINFG---NIDSPYIIKVKISKYDPN-KDDYTTIIQQ 492
QY 172 -LALRSTLYGYSNLIWRSMWNEVAFNNGSGDGDIDKPVVPEQ 216
Db 493 TVTMQTITNEYTGE--FRTASYDNTIAFSTSSGQGG--DLP--PEK 533
```


US-09-200-650E-5

Query Match	15.6%	Score 180.5	DB 4	Length 1315	
Best Local Similarity	27.7%	Pred. No. 1.4e-08			
Matches	59	Conservative 35	Mismatches 100	Indels 19	Gaps 8
QY	7	YIDPENVKKTGN-VTLATGISTTANKTVLDYKFKYFNLSIKG----	----	TIDQDKTNN	61
Db	361	YMDADTTIPVSKNDVEENVITIGNTTKTITANIQYDPYVVNKKNSIGSAFTETVSHVGNKEN	420		
QY	62	T--YRQTIYVNPSSGDNVAPVLTNLKNKENTUSNAL--IDQNTSIKYKVKVDNAADLSSEY	117		
Db	421	PGYVKQTIYVNPSESLTNAKLKQVAYHSSYPNNIGQINKVDTDIKIYQVPKGYTLNKGY	480		
QY	118	FVNPENFEDVNSV--NITEFENQYKVEFNTPPDQITPIYVVVNGHID--PNSKGDIAL	174		
Db	481	DVNTKELTVNQYLQKLYIGDNNNSAVIDFGNAD----	----	SAYVMVNTFKQYTNSESPTLV	536
QY	175	RSTLYGYSNIIWRSMWDNEVAFNNGSGSDG	207		
Db	537	OMATISSTGN--KSVSTGNALGFTNNOSGAG	566		

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

Query Match	15.2k;	Score	175.5;	DB 4;	Length	1027;
Best Local Similarity	19.6k;	Fred. No.	2.9e-08;			
Matches	44;	Conservative	61;	Mismatches	97;	Indels
					23;	Gaps
					97	
Qy	2	LTPAYIDPENVKKTGNVTIATGISTANKTIVLADY-EKYGKEYNLSIKGTIDQIDKTN	60			
Db	310	LEINLFIDPKVTQNGNQTTITSLNEBQTSKELDKYKDGIGNYY-ANLNGSIETFNKAN	368			
Qy	61	NTYRQTIYVNPSSGDNVIAPVLGNLKPNTOSNALIDQONTSIKYK-VDNAADLSSEYFV	119			
Db	369	NRESHVAFIKPNNGKTTSVTVGTLMKSGNQNG---NQPKVRIEYFVLGNNEDEAKSVYA	424			
Qy	120	N---PENFEDVNTSV--NITFEPNQYKVEFNTEDDQITPTTVVWNGHIDPNSKGDIALL	174			
Db	425	NTTDTSKFEKVTSNMGLNLQNGYSLNLI---ENLIDKTYVVHYDGEY-LNGTDEVDVF	479			
Qy	175	RSTLYGYNISII-----WRSMSDNEVAFNNGSGSGDIDKPVV	213			
Db	480	PTOMVGHPEQIYKYVYDRGVTYITWDNGILVYSNKANENGKNGPII	524			

```

RESULT 15
US-08-781-986A-5254
; Sequence 5254, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PE248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5254:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1027 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-781-986A-5254

```

```

Query Match      15.2%; Score 175.5; DB 4; Length 1027;
Best Local Similarity 19.6%; Pred. No. 2.9e-08;
Matches 44; Conservative 61; Mismatches 97; Indels 23; Gaps 9

QY 2 LMPAYIDPENVKKTGNVTLATGIGSTANKVTLVDY-EKYGFYNLSIKGTIDQDKTN 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 310 LEINLFDPKTVQTNGNQITITLNEEQTSKLDVKYKDGIGNY-ANLNGSIETFNKAN 368

```

US-08-956-171E-5254


```

Qy 61 NTRYQTIYVNPESGDNVIAPIVLTGNLKPNTDSNALIDQQNTSIKVKYK-VDNAADLSESFV 119
Db 369 NRFSHVAFIKENNGKTTSTVTGTLMKGSNQNG---NQPKVRIPEYLGNNEDIAKSVYA 424
Qy 120 N---PENFEDYTVSV--NITFPNPNQYKVEFNTPPDDQITTPYIVVVNCHIDPNSKGDIAL 174
Db 425 NITDTSKPFKVTSMNGNLNLQNNGSYSLNI-----ENLDKTYVVHYDGEY-LNGTDEVDF 479
Qy 175 RSTLYGYNSNII-----WRSMWDNEVAFNNGSGSGDGIDKPVV 213
Db 480 RTQWVGHPQYLYKYDYDRGYTLTWDNGLVLYSNKANGNEKNGPII 524

```

Search completed: November 3, 2004, 19:11:52
Job time : 11.7401 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:52:30 ; Search time 6.41939 Seconds
(without alignments)
3282.472 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_550

Perfect score: 1157

Sequence: 1 TLTPAYIDPENVKKTGNVT.....NGSGSGDIDKPVVPEQDE 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1157	100.0	933	2 S41539	fibrinogen-binding
2	991	85.7	989	2 D89852	fibrinogen-binding
3	255	22.0	940	2 S19702	fibrinogen-binding
4	241.5	20.9	1092	2 T30214	fibrinogen-binding
5	237.5	20.5	961	2 G90053	hypothetical prote
6	208	18.0	877	2 F90070	clumping factor B
7	207.5	17.9	1166	2 T28680	fibrinogen-binding
8	200.5	17.3	1141	2 E89824	hypothetical prote
9	180.5	15.6	1315	2 T28679	fibrinogen-binding
10	175.5	15.2	1018	2 A32192	hypothetical prote
11	160.5	13.9	1038	2 H90053	hypothetical prote
12	160.5	13.9	1385	2 D89824	hypothetical prote
13	122.5	10.6	953	2 C89824	hypothetical prote
14	110	9.5	1039	2 T30856	protein F2 - Stre
15	109.5	9.5	1386	2 AC1333	surface protein (L
16	103	8.9	421	2 AC1341	hypothetical prote
17	102.5	8.9	3283	2 AC1018	large repetitive p
18	102	8.8	234	2 T05604	hypothetical prote
19	102	8.8	5005	2 F82884	hypothetical prote
20	101.5	8.8	655	2 A57681	hypothetical prote
21	101	8.7	642	2 D81401	probable flagellar
22	101	8.7	4152	2 T31102	filamentous hemag
23	99	8.6	439	2 AE1251	probable peptidogl
24	98	8.5	691	2 B75622	hypothetical prote
25	98	8.5	1457	2 D81019	adhesion and penet
26	98	8.5	2269	2 T18472	hypothetical prote
27	97.5	8.4	463	2 A44808	cellulase (EC 3.2
28	97	8.4	676	2 C97775	acylamino-acid-rel
29	97	8.4	749	2 S64780	hypothetical prote

30 97 8.4 1711 2 C71625 variant-specific s
31 96.5 8.3 762 2 A34355 cell surface prote
32 96.5 8.3 784 2 AC1091 5'-nucleotidase, p
33 96.5 8.3 1993 2 AF1450 probable peptidogl
34 95.5 8.3 699 2 T18426 hypothetical prote
35 95.5 8.3 995 2 S50358 hypothetical prote
36 95.5 8.3 1487 2 AG2560 hypothetical prote
37 95 8.2 294 2 E82879 ABC transporter UU
38 95 8.2 941 2 F97353 uncharacterized co
39 94.5 8.2 1449 2 B81963 IGA-specific serin
40 94.5 8.2 2539 2 B71619 hypothetical prote
41 94 8.1 648 2 S10869 enterotoxin A - Cl
42 93.5 8.1 526 2 T14875 hypothetical prote
43 93.5 8.1 785 2 AB1455 5'-nucleotidase, p
44 93.5 8.1 1417 2 AG2137 hypothetical prote
45 93 8.0 319 2 F89887 Alpha-Hemolysin pr

ALIGNMENTS

RESULT 1

S41539

fibrinogen-binding protein - Staphylococcus aureus

N;Alternate names: clumping factor

C;Species: Staphylococcus aureus

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S41539; S36630

R;McDevitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.

Mol. Microbiol. 11, 237-248, 1994

A;Title: Molecular characterization of the clumping factor, (fibrinogen receptor) of Staphylococcus aureus

A;Reference number: S41539; PMID:8170386

A;Accession: S41539

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-933 <MCD>

A;Cross-references: UNIPROT:Q53653; EMBL:Z18852; NID:9397525; PIDN:CAA79304.1; PID:939752

Query Match 100.0%; Score 1157; DB 2; Length 933;
Best Local Similarity 100.0%; Pred. No. 3.9e-77;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVTATGISTGTTANKVLTVDYKGFYNLSIKGTIDQIDKTN 60
Db 332 TLTPAYIDPENVKKTGNVTATGISTGTTANKVLTVDYKGFYNLSIKGTIDQIDKTN 391
QY 61 NTRYQTIYVNPFGDNVIAPVLTGNLKPNTDSNALIDQNTSIKVKYVDNADLSESVFN 120
Db 392 NTRYQTIYVNPFGDNVIAPVLTGNLKPNTDSNALIDQNTSIKVKYVDNADLSESVFN 451
QY 121 PENFEDVTSNITFPNPQYKVEFNTPDQITTPYVVVNGHIDPNKGDALRSTLYG 180
Db 452 PENFEDVTSNITFPNPQYKVEFNTPDQITTPYVVVNGHIDPNKGDALRSTLYG 511
QY 181 YNSNIIWRSMWDNEVAFNNGSGSGDIDKPVVPEQDE 219
Db 512 YNSNIIWRSMWDNEVAFNNGSGSGDIDKPVVPEQDE 550

RESULT 2

D89852

fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (strain

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C;Accession: D89852

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizurani, U.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; H

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.

A;Reference number: A89759; PMID:21311952; PMID:11418146

A;Accession: D89852

A;Status: preliminary

A:Molecule type: DNA
A:Residues: 1-989 <KUR>
A:Cross-references: UNIPROT:Q99VJ4; GB:BA000018; PID:gl3700678; PIDN:BAB41975.1; GSPDB:C
A:Experimental source: strain N315
C:Genetics:
A:Gene: Clfa

Query Match 85.7%; Score 991; DB 2; Length 989;
Best Local Similarity 84.9%; Pred. No. 6.6e-65;
Matches 185; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

QY 2 LTMPAYIDPENVKKTGNVTLATGISTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 61
DB 333 ITMPAYIDPENVTKGNVTLTGIGTNTASKTVLIDYKGYQFNLSIKGTIDQIDKTN 392
QY 62 TYRQTIYVNPAGDNIAPVLGNLKPNTDSNALIDQOQNTSIKVKYKVDNAADLSYFNP 121
DB 393 TYRQTIYVNPAGDNIAPVLGNLKPNTDSNALIDQOQNTSIKVKYKVDNAADLSYFNP 452
QY 122 ENFEDVTSNMITFPNPNQYKVEFTPDQITTPYVVVNGHIDPNSKGLALRSTLYG 181
DB 453 SDFEDVTNQVRISFPNPNQYKVEFTPDQITTPYVVVNGHIDPNSKGLALRSTLYG 512
QY 182 NSNIWRSMWNEVAFNNGSGGIDKPKVPVPEQDE 219
DB 513 DSNFTWRSMWNEVAFNNGSGGIDKPKVPVPEQDE 550

RESULT 3
S19702
fibronectin-binding protein B - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 09-Jul-2004
R:Accession: S19702
R:Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibronectin binding proteins in Staphylococcus aureus
A:Reference number: S19702; MUID:92111475; PMID:1837266
A:Accession: S19702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <JOE>
A:Cross-references: UNIPROT:Q53682; EMBL:X62992; NID:g49040; PIDN:CAA44726.1; PID:g58156
C:Keywords: fibronectin binding

Query Match 22.0%; Score 255; DB 2; Length 940;
Best Local Similarity 26.9%; Pred. No. 7.1e-11;
Matches 61; Conservative 54; Mismatches 82; Indels 30; Gaps 8;

QY 2 LTMPAYIDPENVKKTGNVTLATGISTTANKTVLVDY----EKYKGFNLSIKGTIDQI 56
DB 273 LSNLFIPTTPTVKGNQNVKLGTTVSKLFNIOYLGVRDNGW----VTANGRIDTL 328
QY 57 DKNNTYRQTIYVNPAGDNIAPVLGNL----KPNTDSNALIDQOQNTSIKVKYKVDNAAD 112
DB 329 NKVDGKSFHFAYMKPNQSLSSVTVTGQVTKGNKPGV-----NPTVKVYKHIGSD 380
QY 113 LSESYFNPEN---FEDVTSNMITFPNPNQYKVEFTPDQITTPYVVVNGHIDPNSK 169
DB 381 LAESVYAKLDVSKFEDVTNMSLDFDTNGGYSFNFNLDQ---SKNYVIKGYGYDSNA- 437
QY 170 GDALALRSTLYGNSNIWRSMWNEVAFNNGSGGIDK----PVV 213
DB 438 SNLEFQTHLFGYYNYTSLTWNGVAFYSNNAQGDGKLEPII 484

RESULT 4
T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30214
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66, 2666-2673, 1998
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.
A:Reference number: Z20781; MUID:98261511; PMID:9596732
A:Accession: T30214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1092 <NIL>
A:Cross-references: UNIPROT:O70022; EMBL:Y17116; NID:el296734; PID:el296735; PIDN:CAA766

Query Match 20.9%; Score 241.5; DB 2; Length 1092;
Best Local Similarity 31.4%; Pred. No. 8.5e-10;
Matches 71; Conservative 37; Mismatches 93; Indels 25; Gaps 10;

QY 2 LTMPAYIDPENVKKTG---NVTLATGISTTANKTVLVDYKGYKFNLSIKGTIDQIDK 58
DB 389 LKLSYIDKSKVPNNNTKLDVEYKALSS--VNKTIITVEYQRPNEKRTANLQSMFTNIDT 446
QY 59 TNNYRQTIYVNPAGDNIAPVLGNLKPNTDSNALIDQOQNTSIKVKYKVDNAADLSYF 118
DB 447 KNHTVEQTIYINPL--RYSAKETNNVNSGDEGTIDDTIIIKVYKVGQNQLPDSNR 504
QY 119 V-NPENPEDVTSNMITFPNPNQYKVEFTPDQITTPYVVVNGHIDPNSKGD----- 171
DB 505 IYDSEYEDVTNDYQALGNNDVYNIFG----NIDSFYIIVKISKYDEN-KDDVTTIQQ 559
QY 172 -LALRSTLYGNSNIWRSMWNEVAFNNGSGGIDKPKVPVPEQ 216
DB 560 TVTWTQTTINEYTG--FRTASYDNTIAFSTSSGQGG-DLP--PEK 600

RESULT 5
G90053
hypothetical protein fnbB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
R:Accession: G90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, R.; Inoue, R.; Kato, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A9758; MUID:21311952; PMID:11418146
A:Accession: G90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-961 <KUR>
A:Cross-references: UNIPROT:Q99RD3; GB:BA000018; PID:gl3702452; PIDN:BAB43593.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnbB

Query Match 20.5%; Score 237.5; DB 2; Length 961;
Best Local Similarity 26.6%; Pred. No. 1.4e-09;
Matches 61; Conservative 47; Mismatches 94; Indels 27; Gaps 7;

QY 2 LTMPAYIDPENVKKTGNVTLATGISTTANKTVLVDY-----EKYKFNLSIKGTIDQI 56
DB 276 LNLNLFIDPTTPTVKQGSQKVEVTLGONKYSKEFDIKYLDGVKDRMG----VTYNGRIDTL 331
QY 57 DKNNTYRQTIYVNPAGDNIAPVLGNLKPNTDSNALIDQOQNTSIKVKYKVDNAADLS 116
DB 332 NKEEGKSFHFAYVKPNQSLTSVTVTGQVTSYGKQSA---NNPTVKVYKHIGSDELAES 387
QY 117 YFV---NPNFEDVTSNMITFPNPNQYKVEFTPDQITTPYVVVNGHIDPNSKGLA 173
DB 388 VYAKLDVTSKFEVDTEKVNLSYTSNGGYTLNLGLDN--SKDYVIKGYEYDQNAK-DLN 444
QY 174 LRSTLYGNSNIWR-----SMWNEVAFNNGSGGIDKPKVPV 214
DB 445 FRTHLSGYHKYYPYYPYYPYYPVQLTWNGVAFYSNNAKGDGKKNNDP 493

RESULT 6


```
Db 361 YMDADTIPVSKNDVEFNVTIGNTTTKTIANIQPYDVVNEKNSIGSAFTETVSHVGNKEN 420
QY 62 T--YRQTIYVNPNGDNVIAPIVLGTNLTNPNTSNAI--IDQNTSITKVYKVDNAADLSBSY 117
Db 421 PGYKQTIYVNPSENLSNAKLKVOAYHSSYPNNIGQINKDVTDIKIYQPKGYTLNKG 480
QY 118 FVNPENFEDVNSV--NITFPNPQYKVEFNTPDQITTPYIVVNGHID--PNSKGDAL 174
Db 481 DVNTELDTVNOYLQKITYGDNNSAVIDFGNAD----SAYVVMNTKPYQYNSPSPIV 536
QY 175 RSTLYGYSNIIWRSMWDNEVAFNNGSGDGD 207
Db 537 QMATLSSTGN--XSVSTGNALGFTNNQSGGAG 566

RESULT 10
A32192
hypothetical protein - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 16-Feb-1997
C:Accession: A32192
R:Signaes, C.; Kaucii, G.; Joensson, K.; Lindgren, P.E.; Anantharamaiah, G.M.; Hoesek, M
Proc. Natl. Acad. Sci. U.S.A. 86, 699-703, 1989
A:Title: Nucleotide sequence of the gene for a fibronectin-binding protein from Staphylo
A:Reference number: A32192; MUID:89098998; PMID:2521391
A:Accession: A32192
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1018 <Sig>
A:Cross-references: GB:J04151
C:Keywords: fibronectin binding

Query Match 15.2%; Score 175.5; DB 2; Length 1018;
Best Local Similarity 19.6%; Pred. No. 5.5e-05;
Matches 44; Conservative 61; Mismatches 97; Indels 23; Gaps 9;

QY 2 LTMPAYIDPENVKKTGNVTLATIGSTTANKTIVLDY-EYKGFYNLSIKGTIDQIKTN 60
Db 301 LEINLFIDPKTVQTNQITITSLNEEQTSKELDKYKDGIGNY-ANLNGSIETFNKAN 359
QY 61 NYRQTIYVNPNGDNVIAPIVLGTNLTNPNTSNAIIDQNTSITKVYKVDNAADLSBSYFV 119
Db 360 NRFSHVAFIKPNNGKTTSTVTGTLMKGSNQNG---NQPKVRIPEYLGNNEDIAKSVYA 415
QY 120 N---PENEDVTNSV--NITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNSKGDAL 174
Db 416 NTTDSKFKFVTSNMSGNLNLQNGSYSINI----ENLDKTYVHVHVDGEY-LNGTDEVDF 470
QY 175 RSTLYGYSNII-----WRSMWDNEVAFNNGSGDGDIDKPVV 213
Db 471 RTQMGHPQLKYKYDYRGYTLTWNGLVLYSNKANGNEKNGPII 515

RESULT 11
H90053
hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: H90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A9758; MUID:21311952; PMID:11418146
A:Accession: H90053
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1038 <KUR>
A:Cross-references: UNIPROT:Q99RD2; GB:BA000018; PID:g13702453; PIDN:BAB43594.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb
```

```
Query Match 13.9%; Score 160.5; DB 2; Length 1038;
Best Local Similarity 22.8%; Pred. No. 0.00071;
Matches 53; Conservative 49; Mismatches 107; Indels 23; Gaps 10;

QY 2 LTMPAYIDPENVKKTGNVTLATIGSTTANKTIVLDYKGYKFNLSIKGTIDQIKTN 61
Db 295 LEINLFIDPKTVQSNGEQKITSKLNGEETKTIPIVYVNPNGVSNSTYVNGSIETENKSN 354
QY 62 TYRQTIYVNPNGDNVIAPIV-LTGNLKPNTDSNALIDQNTSITKVYKVDNAADLSBSYFV 119
Db 355 KFTIAYIKPMNGSQSNSTVSTGTI---TEGSLAGQPT-VKVYFYLGKKDELQSVYA 410
QY 120 NPEN---PEDVTNSV--NITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNSKGDAL 174
Db 411 NTSOTNKEFKVTKEMNGKLSVQDNGSYSINL----DKLDTYVHVHTGEYLQGS-DQVNF 465
QY 175 RSTLYG-----YNSNIIWRSMWDNEVAFNNGSGDGDIDKPVVPPQDFE 219
Db 466 RTLYGYPERAYKSYVYGGYRLTWNGLVLYSNKADGNGKNGQIIQDNDFE 517
```

```
RESULT 12
D89824
hypothetical protein sdrB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D89824
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A9758; MUID:21311952; PMID:11418146
A:Accession: D89824
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1385 <KUR>
A:Cross-references: UNIPROT:Q99W47; GB:BA000018; PID:g13700454; PIDN:BAB41751.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: sdrB

Query Match 13.9%; Score 160.5; DB 2; Length 1385;
Best Local Similarity 25.1%; Pred. No. 0.001;
Matches 60; Conservative 37; Mismatches 99; Indels 43; Gaps 10;

QY 1 TLTMPAYIDPENVKKTG-----NVLTATIGSTTANKTIVLDYE 39
Db 339 TYTFTDYVDRFNSVRKMGINSIYMDADTIPVDKDVFPFSTIGNQITTTAD----ITYP 394
QY 40 KYGKFYNLSIKG---TIDQIDKTNT--YRQTIYVNPNGDNVI-APVLGTNLTNPNTDSN 92
Db 395 AYKEADNNSIGSAFTETVSHVGNVEDPGYVNVNPMKDLKGAKLKVAYHPKYPIN 454
QY 93 -ALIDQNTSITKVYKVDNAADLSBSYFVNPENFEDVNSV--NITFPNPQYKVEFNTPD 149
Db 455 IQGINQNVNIIKIVRPEGYTLNKGVDVNTDLVDVTEBFKNMVTGVSQSNVLDGFG--- 511
QY 150 DOITTPYIVVNGHID-PNSKGDALRLSTLYGYSNIIWRSMWDNEVAFNNGSGDGD 207
Db 512 -DITSAYVVMVNTKPYQYNSPSPIVQMATLSSTGN---KSVSTGNALGFTNNQSGGAG 566
```

```
RESULT 13
C89824
hypothetical protein sdrC [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C89824
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
```

Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: C89824
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-953 <RUR>
A;Cross-references: UNIPROT:Q99W48; GB:BA000018; PID:gl3700453; PIDN:BA81750.1; GSPDB:G
A;Experimental source: strain N315
C;Genetics:
A;Gene: sdrC

Query Match 10.6%; Score 122.5; DB 2; Length 953;
Best Local Similarity 23.4%; Pred. No. 0.39;
Matches 54; Conservative 32; Mismatches 108; Indels 37; Gaps 8;

```
Qy      1 TLTWPAIDP-ENVK-----KTGNVT-----LATIGSTTANKVLVDYEKVKG 43
Db     281 TYTFNTVNDQYTNNVSGSFEQAFAKRENATTDKYAMEVFVTLNGDTSYKDVIDYGKQG 340
Qy     44 FYNLSIKGTDIQIDKTNTRYQTLYVNPSGDNVIAPLTGNL---KNPTDSNALIDQQNT 100
Db    341 QQLISSNYYINNELSRN---MTVIYVQPCKTYTKETFTVNLTCYKFNPDAK----- 389
Qy   101 SIKYKVDNAADLSSEYFVPNPENPEDVTNSVNIITFPNQ-YKVEFNTDPDQITTPYIV 159
Db   390 NFKIYEVDQNQFVDSFTPDTSKLDVYTGQFDVIYSNDNKTTATVDLLNGOSSDKQYIIQ 449
Qy   160 VNHGHDPNSSKGDLALRSLTLCYNSNIITWRSMWDNEVAFFNNGSSGGDIGDK 210
Db   450 OVAYXPDNSSDTDGKIDYTILETONG----KSWSNYSNVNGSSSTANGDOK 495
```

RESULT 14

T30856 protein F2 - Streptococcus pyogenes

C:Species: Streptococcus pyogenes

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C:Accession: T30856

R:Jaffe, J.; Natanson-Yaron, S.; Caparon, M.G.; Hanski, E.

Mol. Microbiol. 21, 373-384, 1996

A:Title: Protein F2, a novel fibronectin-binding protein from *Streptococcus pyogenes*, p4

A:Reference number: Z20907; MUID:97011581; PMID:8858591

A:Accession: T30856

A:Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: DNA

A:Residues: 1-1039 <JAF>

A:Cross-references: UNIPROT:P72534; EMBL:U031980; NID:G1654115; PID:G1654116; PIDN:AAC445

C:Genetics:

A:Note: PRTF2

Query Match 9.5%; Score 110; DB 2; Length 1039;
Best Local Similarity 21.3%; Pred. No. 3.6;
Matches 49; Conservative 43; Mismatches 98; Indels 40; Gaps 9;

[illegible]

RESULT 15

AC1533
surface protein (LPXTG motif) [imported] - *Listeria innocua* (strain Clip11262)
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AC1533
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kraft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlrand, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1533
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1386 <GIA>
A:Cross-references: UNIPROT_Q92DL0; GB:AL592022; PIN:CAC96035.1; PID:gl6143254; GSPDB:GN
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0803

Query Match 9.5%; Score 109.5; DB 2; Length 1386;
Best Local Similarity 21.0%; Pred. No. 5.8;
Matches 57; Conservative 29; Mismatches 68; Indels 117; Gaps 10;

QY 8 IDPENVKKTGNVTLAT-----GIGSTTANKTVLDVY----- 38

Db	1066	IDNSVKVTGSVDIQTPCKYRITYSIEG-ASTHTVTVLDAQSNLVAKNSTIYVGDKWCS	1124
QV	39	-----EKYGFYNLS---IKGTIDQIDKTNNTRYQTIYVN-----	70

1125 KONFVSATDKYKPIDLSLLTIVTGTVD--TTTPEYEITYSVNGLTTTITVTVKENQANI 1182

QY	71	-----PSGDNIAPV-----	-----LTGNLKPNDSN-----	92
----	----	---------------------	-----------------------	----

Db 1183 VAEDSTTHAKESKKAADNFVSATDKTGNIGLSSVSVTGKVDINTPGNYEITYTIDGVST 1242

QY 93 -----ALIDQONTSIKV---YKVDNAADLSESYFVNPFENFEDVTVNSVNIITFPNPN 139

Db 1243 TITVTLVNHSQIEAHDAKIKVGTSSWKPEDSFILAKDKFGETADFSDVTVEGTVDTTPG 1302

Qy 140 QYKVFENTDDQITTPYIVVNGHIDPSKG 170

Db 1303 KYQITYTIDGVSVTITVIV-----EDNSKG 1327

Search completed: November 3, 2004, 19:04:07
Job time : 8.41939 secs

This Page Blank (uspto)